PROPERTIES OF MATTER

PROPERTIES OF MATTER

BY
H. N. V. TEMPERLEY, M.A.
Fellow of King's College, Cambridge
Sometime Smithson Research Fellow of the Royal Society
Lecturer in Physics in the University of Cambridge

LONDON
UNIVERSITY TUTORIAL PRESS LTD
Clifton House, Euston Road, N.W. 1

All rights reserved. No portion of the book may be reproduced by any process without written permission from the publishers.

8A-282

JAMMU & KASHMIR UNIVERSITY
LIBRARY
No. 12031
Date 19-5-55
SRINAGAR

CHECKED

Catalogued

Published 1953

530.1
T246R

ALLAMA IQBAL LIBRARY
12031

PRINTED IN GREAT BRITAIN BY UNIVERSITY TUTORIAL PRESS LTD, FOXTON, NEAR CAMBRIDGE

PREFACE

This book is the result of eight years' experience in teaching physics at Cambridge, and is primarily intended for students preparing for the Natural Sciences and Mechanical Sciences Triposes at Cambridge, the B.Sc. General and Special Examinations at London, and similar examinations elsewhere. An attempt has also been made to cater for research workers who design their own apparatus and for those who have a fundamental interest in any of the fields treated. A typical "precision" measurement (of the acceleration of gravity) is described in detail, and brief introductory treatments of a number of advanced topics, with references to specialised treatises or original papers, have also been included.

The examples at the end of each chapter are intended to illustrate and amplify portions of the text, and hints on the more difficult ones are included. The examination questions are a representative selection from recent London and Cambridge papers, and thanks are due to the Senate of the University of London and to the Syndics of the Cambridge University Press for permission to include them. To help the student gain experience in tackling questions under "examination conditions", hints on these questions have been relegated to the "Answers" section.

The author wishes to thank Dr M. N. Hill, of King's College, Cambridge, for help with the chapter on gravitation.

The author, however, is solely responsible for any expressions of opinion and, if the book contains errors or misprints, he would be grateful to receive notice of them.

H. N. V. T.

King's College, Cambridge.

CONTENTS

CHAPTER I. HISTORICAL INTRODUCTION 1

Continuum and atomic theories of matter. Mechanics and gravitation. Deformation of solids, hydrodynamics, surface tension, properties of solutions, and allied phenomena

CHAPTER II. PRINCIPLES OF MECHANICS... 7

Laws of motion. Mass and force. Conservation laws. Virtual work. Particle dynamics. Orbits. Rigid body dynamics. Compound pendulum. Centre of gravity and moments of inertia. Instantaneous axis of rotation. Gyroscopic motion.

CHAPTER III. GRAVITATION 45

Newton's Law. Gravitational field and potential. Inertial and gravitational mass. Measurements of the gravitational constant *G*. Measurement of *g*, the acceleration of gravity and its variations from place to place. Gravity surveys and their applications. Joly's theory of the rise and fall of the continents. Einstein's theory of gravitation and the evidence for it.

CHAPTER IV. THE DEFORMATION OF SOLIDS. ELASTICITY AND PLASTICITY 72

Dilatation and shear. Stresses in a solid and relation with deformations. Two general theorems on elasticity. Elastic constants and relations between them. Stress calculations. Bending moment and shearing force, relations with deformation. Measurement of elastic constants. Variation with temperature. Elastic stability. Whirling of shafts. Elastic waves.

CHAPTER V. KINETIC THEORY OF GASES 113

Assumptions made. Maxwell's distribution law. Equipartition of energy. Boltzmann's expression. Sedimentation. Brownian movement. Mean free path. Transport properties. Viscosity, thermal conductivity, thermal transpiration, thermal diffusion, true diffusion. Diffusion pumps. Phenomena at very low pressures. "Knudsen gas." Slip and accommodation coefficients. Measurement of low pressures.

CHAPTER VI. HYDRODYNAMICS 138

Definition of liquid. Expansion and shear viscosities. Equations of motion. Solutions in some simple cases. Types of flow in a tube: Poiseuille, Knudsen, and turbulent. Conditions behind an obstacle. Irrotational motion. Effect of boundaries. Simple cases of irrotational motion. Rayleigh disc. Waves on water ripples and deep-water waves. Tidal waves. The tides. Non-linear hydrodynamics. Shock fronts. Hydraulic jump.

CHAPTER VII. THE PHYSICAL PROPERTIES OF LIQUIDS ... 182

Observed properties. Various models of a liquid, "gas-like" and "solid-like". Van der Waals' "cell", "hole", rigid sphere, and Debye models and their successes. Superheating of liquids. Measurement of the properties of a liquid. Bulk modulus, tensile strength, viscosity. Special properties of liquid helium. Ultrasonic absorption in liquids and its interpretation.

CHAPTER VIII. SOLUTIONS AND ELECTROLYTES. OSMOTIC PRESSURE, DIFFUSION, AND RELATED PHENOMENA ... 210

Definition and measurement of osmotic pressure. Relationship with vapour pressure, boiling point, and freezing point. Theory of osmotic pressure. Solubilities. Electrolytes. Weak and strong electrolytes. Debye-Hückel theory. Diffusion in liquids.

CHAPTER IX. SURFACE TENSION, ADSORPTION, AND ALLIED EFFECTS 231

Surface tension and surface energy. Angles of contact. Pressure difference across a curved liquid surface. Stability conditions for soap films. Form of meniscus. Effect of curved surface on vapour pressure. Forces between floating bodies. Measurement of surface tension. Spreading of liquids. Surface tension of solutions. Laplace theory of surface tension and its relation to other properties. Variation of surface tension with temperature. The parachor. Association in a liquid. Surface tension of solids. Adsorption. Friction and lubrication.

CHAPTER X. DIMENSIONAL METHODS. DYNAMIC SIMILARITY 276

Method of dimensional analysis. Initial assumption and conclusions compatible with it. Model experiments. Dynamic similarity. Resistance of ships. Checking of physical equations.

MATHEMATICAL APPENDIX 285

The damped simple harmonic oscillator. Free and forced oscillations. Resonance. Effects of non-linearity. Fourier analysis. Full and half range series. Application to wave and diffusion equations. Integrals occurring in the kinetic theory of gases.

EXAMINATION QUESTIONS 303

TABLES 314

ANSWERS TO EXAMPLES 316

HINTS AND ANSWERS TO EXAMINATION QUESTIONS 317

INDEX 319

PROPERTIES OF MATTER

CHAPTER I

HISTORICAL INTRODUCTION

1. The continuum theory of matter

Men have speculated about the structure of matter, and have been aware of such properties as weight, cohesion, and elasticity, ever since they have been capable of thought, so that our subject may fairly claim to be one of the oldest branches of science. The first attempts to interpret such properties as elasticity were based on the assumption that matter is *continuous* or capable of infinite division, but theory could not make very much progress until the invention of the calculus by **Newton** and **Leibnitz**. Although the continuum theory has long been known to be incorrect, there being overwhelming evidence for the *atomic* structure of matter, the concept of continuous matter is still often helpful on account of the extreme smallness of atoms. The concept is used freely in this book, *e.g.* in the study of elasticity and hydrodynamics, on the understanding that the predictions of such theories must be regarded as suspect whenever they involve lengths which are not large compared with interatomic distances.

2. The atomic theory of matter

The first definite formulation of an atomic hypothesis seems to be due to the Greek scientist **Leucohippus**, but little use seems to have been made of it, owing to the absence of any experimental evidence, until **Dalton** realised its great value in interpreting the chemical laws of definite and multiple proportions. Applications of the hypothesis to the explanation of the physical properties of gases has proved extraordinarily fruitful, and the pioneer work of **Maxwell** and **Clausius** is being carried on at the present day by such men as **Chapman**, **Knudsen**, and **Lennard-Jones**. Considerable progress has also been made in explaining the properties of solids by means of the kinetic theory of matter, but it must be said that we are still a long way from the ultimate objective of being able to predict all the macroscopic physical properties of a substance if the nature of the interactions between the molecules is assumed to be known. No difficulties of principle seem to remain save the purely mathematical ones involved in treating problems involving large numbers of variables. So far, only theories postulating interaction functions of fairly simple types can be carried through to the point of making firm numerical predictions of properties of solids and gases, while application of such theories to liquids is still in an early stage.

For this reason, the continuum hypothesis, though essentially incorrect, is still extremely useful for descriptive purposes. Some of the consequences of the discrete structure of matter can be allowed for even within the framework of a continuum theory; for example, one can introduce the concept of *viscosity* into hydrodynamics without specifying the precise mechanism by which it arises. Any such descriptive theory may break down under extreme conditions, for example, the concept of viscosity is no longer of value if we are dealing with gases at very low pressures, when mean free paths are becoming comparable with the dimensions of the apparatus (p. 130).

3. Mechanics

No progress in either the atomic or the continuum theories of matter would have been possible unless the laws of mechanics had first been worked out. Concepts such as velocity, force, and inertia, have long existed, but the founder of mechanics was probably **Galileo,** and the explicit statement of the laws of mechanics is due to **Newton.** Newton was principally concerned with the laws of motion of a structureless particle, but such men as **Gauss, Lagrange, Laplace,** and **Hamilton,** developed dynamical theory to a point at which it could be directly applied to systems of many particles with interactions, without the mathematics becoming impossibly complicated. The advent of the theory of relativity and the quantum theory has shown that the laws of classical mechanics are not quite correct, and we now have a clear idea of the precise circumstances in which they may be expected to fail. Speaking roughly, such failures become noticeable for high velocities and for very small, or very large, masses and distances. We shall meet only a few instances of such failures in this book (pp. 68, 202), and, for all other purposes, we shall use Newton's laws, which are usually extremely close approximations to the truth.

4. Gravitation

The concept of *weight* must have been vaguely familiar to man for thousands of years, but no one before **Newton** seems to have even suspected the essential connection between the weight of a terrestrial body and the motion of the moon and planets, as embodied in the law of universal gravitation. Even if anyone did previously have such a suspicion, he could not have anticipated Newton and verified it quantitatively, in the absence of a systematic theory of dynamics, which, in turn, had had to wait on the invention of the calculus. Nothing has occurred since Newton's time to shake in any way our belief in the essential soundness of his conception of a law of gravitation applying universally to all matter; the only modification required by the theory of relativity has been to show that the force between two masses does not always vary precisely as the inverse square of their distance, as Newton supposed (p. 45).

5. Elasticity and plasticity

The systematic study of elastic phenomena dates from the enunciation by **Hooke** of his famous law, but real progress did not come until **Poisson's** realisation that deformations in different directions could not, in a real substance, be treated as independent of one another, soon after which **Lamé** made the first crude attempt at a molecular theory of elasticity, while **Fresnel, Love, St Venant,** and others, showed that many idealised problems in elasticity could be solved exactly by building up a theory quite analogous to electrostatics and electromagnetics. A further fundamental advance was made by **Coker,** who introduced the technique of *photo-elasticity*, that is of stressing transparent materials and examining them optically, which enables one to solve problems in elasticity too involved for direct computation. All these lines of approach have been followed up by workers too numerous to mention, but in recent years attention has been concentrated more on the behaviour of solids when stressed beyond the elastic region. For certain metals, particularly iron and steel, there is a well-defined region of *plasticity*, that is to say a region in which deformation increases steadily with time without the load being further increased. In recent years much attention has been paid to this region by metallurgists, engineers, and physicists, in an attempt to gain insight into the reasons for the failure of metals under stress, and to relate the properties of large pieces of metal to the properties of metallic single crystals, and these in turn to the forces binding the atoms together.

6. Hydrodynamics

The foundation of hydrodynamics is usually ascribed to **Daniel Bernoulli**, but the honour is also shared by **Lagrange** and **Euler.** For a long time the subject developed as a branch of applied mathematics because so many of the problems that could be handled mathematically were too idealised to have practical applications. The subject was not securely linked up with physics until the time of **Lord Kelvin, Stokes,** and **Lamb.** The reason for this was that so many of the results predicted for a fluid without viscosity are never even approximately realised in actual fluids, while attempts to take viscosity into account made all but the very simplest problems mathematically insoluble. In practice it has been found that the hydrodynamic conditions near a solid-liquid boundary are very far from simple, and **Prandtl** founded the *boundary-layer theory*, which is an intensive study of conditions in this region, in an attempt to find out what determines the resistance experienced by a solid when it is moving through a liquid or gas. A second line of approach to this problem has been the direct study of the eddies or *turbulence* found behind a solid obstacle, work begun by **Reynolds,** and being carried on at the present day by **G. I. Taylor** and others.

7. Surface tension

Many of the facts of surface tension must have been known for thousands of years, but the introduction of precise experimental methods for the study of surface phenomena is probably due to **Quincke** and **Lord Rayleigh.** The first work on the interpretation of surface phenomena in terms of intermolecular forces was due to **Laplace,** whose ideas were subsequently developed by **van der Waals.** A further advance was the introduction by **Sugden** of the *parachor*, a quantity which depends on surface tension but is constant over fairly wide ranges of temperature (p. 264). This quantity is believed to be a rough measure of the effective *volume* of molecules, and can be used in helping to decide between various possible chemical structures, though this method is now almost superseded by the much more precise information that is available from X-ray data and infra-red spectroscopy.

In recent years the subjects of surface chemistry and surface physics have assumed considerable importance. The role played by surface phenomena in the action of chemical catalysts had long been suspected, but this subject is only just beginning to be put on a precise basis. It has also been found that surface films can, in certain circumstances, behave as two-dimensional counterparts of gases, liquids, or solids, according to the size of the area they occupy. Much of the early work on surface phenomena has proved to be of little value because the necessity of working under absolutely "clean" conditions was not realised.

8. Theory of solutions

It must be admitted that we are still far from understanding the reasons for the widely differing solubilities of substances that one might expect to be fairly alike chemically, for example, barium sulphate and zinc sulphate, or the various allotropic forms of sulphur. Indeed, it can be said that comparatively little is known about *concentrated* solutions, but very considerable advances have taken place in the theory of *dilute* solutions. These advances fall under two heads, there is first the working out of the relationships between osmotic pressure and the relative lowering of vapour pressure, which can, in turn, be related to the changes of boiling point and freezing point of the solution as the concentration is changed (see p. 211). This branch of the theory of solutions was founded by **van't Hoff.** The other important advance is due to **Arrhenius,** who introduced the *ionic theory*, namely, the postulate that charged ions exist in solution even when electrolysis is not taking place (p. 220). Very many of the properties of electrolytic solutions can be interpreted by means of this idea, and a further advance was made by the **Debye-Hückel** theory, which attempts to take account of the electric interaction of the ions. This latter theory, though fairly successful in accounting for the facts, has been shown to be subject to serious logical difficulties, and later work in this particular field has consisted largely of attempts to consolidate ground that was at one time thought to have been securely won. Very recently, similar theories have been applied to *colloidal* solutions, whose stability depends upon the colloidal particles being electrically charged, but there are serious mathematical difficulties in working out such theories in an absolutely consistent way.

9. Conclusion

It may fairly be said that practically all of the subject-matter of this book is included in the domain of "classical" physics and, with one or two exceptions, does not need the results either of relativity theory or of quantum theory. It may well be that the spectacular advances of quantum theory and relativity have lured away some workers from this older field, within which no especially striking advance has taken place in the twentieth century, but there is no sign of any slowing-down of progress, quite the reverse, and the subject is still a very long way from being a "closed" one. Perhaps it will one day become possible to predict all physical properties of a substance in terms of the molecular interaction function, or vice versa, but that stage is not yet in sight.

The relationship between mathematics and this branch of physics remains almost exactly as it was in the nineteenth century, the object of theoretical work being both to predict numerical results and to give a physically satisfying "picture" of what is taking place. For reasons that cannot be discussed here, the mathematics of relativity and quantum theory has tended to become a formalised skeleton which it is becoming increasingly difficult to clothe with physical flesh. The situation has been ably summed up by **Dirac** thus: "The only object of theoretical physics is to calculate results that can be compared with experiment, and it is quite unnecessary that any satisfying description of the phenomena should be given". This defeatist attitude may be a bitter necessity in quantum theory, but the author would strongly deprecate any tendency to assume such an attitude in this branch of physics, where one is still entitled, and indeed compelled, to insist that no concepts be introduced that cannot be precisely defined both in a mathematical sense, and in the sense that clear "physical pictures" of them can be formed.

BIBLIOGRAPHY FOR CHAPTER I

W. C. Dampier, *A History of Science.* 4th Edn. C.U.P. (1948).

J. Arthur Thomson (ed.), *The Outline of Science.* Eyre and Spottiswoode, London (1922).

CHAPTER II

PRINCIPLES OF MECHANICS

1. Newton's laws of motion

The whole of mechanics can be derived from Newton's laws of motion, and it has been thought advisable to include such derivations of some of the more important results from these laws, in order to illustrate the fundamental principles used, despite the fact that most of the users of this book will already be familiar with elementary statics and dynamics.

Newton's First Law.—A body persists in a state of rest or uniform motion in a straight line, unless it is subjected to an external force.

Newton's Second Law.—For a given body, the rate of change of momentum is proportional to the external force acting upon it, and the change in momentum takes place along the line of action of that force.

Newton's Third Law.—If a body A exerts a force on a body B, then body B exerts an equal and opposite force on body A.

It is extremely difficult, as always in science, to avoid circularity in elementary definitions, unless we are prepared to treat some of our concepts as *indefinable*, or unless we define the qualitative ideas by direct appeal to the senses, afterwards specifying means of measurement. No real difficulty arises with the concepts of displacement and velocity, but attempts to define force or mass usually end as paraphrases of the first two laws of motion. We shall regard *force* as an indefinable, appealing to our senses for a rough *description*, and it may be *measured* either by the extension it produces on a spring-balance, or by the acceleration produced when it acts on a standard test-body. In terms of force we can define the *inertial mass* of a body by means of Newton's second law, as the constant of proportionality between the force acting on that body and the acceleration this force produces. A definition of *gravitational mass* will be given later (p. 49).

The *centimetre* is defined in terms of the standard metre kept at Sèvres, the *second* in terms of the mean time of rotation of the earth relative to the sun (the mean solar day). Since it is not easy to make an accurately permanent standard of force, it is better to use mass as our third standard, the gram is thus defined in terms of the inertial mass of the standard kilogram. The unit of force, the dyne, is that force that gives one gram an acceleration of one centimetre per second per second.

Thus if the mass of a body is M, and g is the acceleration due to gravity, the force due to the weight of a body is Mg dynes.

Newton's laws are difficult to verify by direct experiment because of the disturbing effects of friction. The most accurate method is indirect, the laws of dynamics are deduced from them and applied to astronomical problems and the predictions of dynamics are compared with what is actually observed, the agreement being close enough to enable predictions to be made years ahead. There are certain types of force occurring in nature for which Newton's third law is untrue, for example, the forces between *moving* electric charges. However, a very large number of forces are *central*, that is the forces between any two particles act directly along the line joining them, and the third law always seems to hold for such forces.

2. The parallelogram law

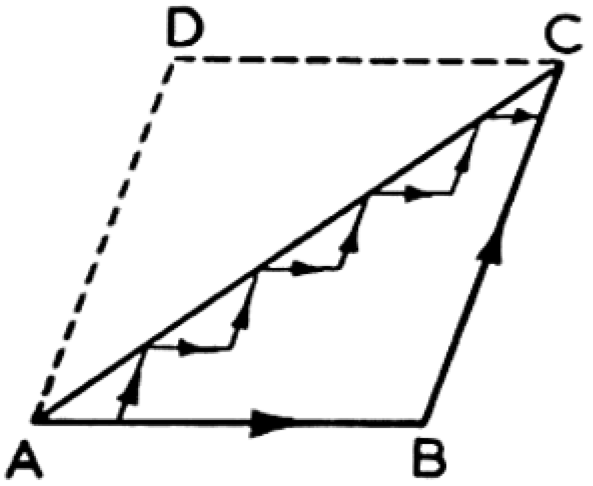


FIG. 1.

It is not self-evident that *any* quantity that is a *vector*, having *direction* as well as *magnitude*, is subject to the parallelogram law of addition. Vector quantities that do obey this law are called *polar* vectors, those that do not are called *axial* vectors. A simple example of an axial vector is angular velocity, to which the parallelogram addition law applies only in special cases (see p. 36).

It is therefore necessary to give a formal proof that forces can be added together by the parallelogram law, and this can be done quite easily by means of the second law of motion. Fig. 1 illustrates the self-evident fact that two *successive* displacements AB and BC result in the displacement AC. The problem is to determine the displacement if a point is displaced *simultaneously* in the directions AB and BC. The ziz-zag line between A and C represents the effect of a displacement amounting to $\frac{AB}{n}$ parallel to AB, followed by a displacement of $\frac{BC}{n}$ parallel to BC, then by another displacement of $\frac{AB}{n}$ parallel to AB, and so on. As n becomes very large, this process becomes equivalent to applying two *simultaneous* displacements in the directions AB and AD, the latter being parallel to BC, so that the parallelogram law applies to simultaneous displacements. If we now suppose that these displacements take place uniformly in time we infer that constant velocities obey the parallelogram law as well as displacements. If the two velocities to be compounded are not constant in time, we can still prove the parallelogram law for *velocities* by taking very short time intervals, in which case the parallelogram $ABCD$ would become very small, but the two velocities would be nearly constant over each time interval. If velocities obey the parallelogram law, it follows in the same way that any two *changes of velocity* that take place in the same short time interval can also be compounded by the parallelogram law, that is that *accelerations* obey the law. Since, for a given body, acceleration is proportional to force and lies in the same direction, it follows that *forces* obey the parallelogram law.

3. The conservation laws

Newton's laws of motion lead directly to four very important principles which will be derived below.

(a) *The conservation of momentum*

An assembly of bodies, subject to no external forces, has a total momentum, which is represented by a vector that does not change with time.

By the first law, the momentum of any body can only be changed by the application of a force; whereas, by the third law, the forces exerted on one another by two bodies are always equal and opposite, whether they arise from collision, or from distant interactions of the bodies through, *e.g.* gravitational attraction. Since, by the second law, rate of change of momentum is proportional to force, the momentum lost by one body in a time dt is equal to $p\,dt$, where p is the force acting on it, and this is exactly balanced by a momentum $p\,dt$ gained by the body with which it is interacting. One may conclude further that if an external force (due to the effect of bodies outside the assembly) acts on the assembly, the rate of change of momentum of the whole assembly is proportional to this force, any forces acting *between* two of the bodies in the assembly being unable to affect their *total* momentum in any direction.

(b) *The conservation of energy*

An assembly of bodies, subject to no external forces, has a constant total energy.

By Newton's second law we have, if a force $\mathbf{F}$, having components F_x, F_y, F_z, acts on one of the bodies of mass m_1,

$$F_x = \frac{d}{dt}(m_1 v_{x(1)}) = m_1 \frac{dv_{x(1)}}{dt} \text{ if } m_1 \text{ is constant.}$$

Multiplying by $\dfrac{dx_1}{dt} = v_{x(1)}$, the velocity of the mass m_1 in the x-direction

$$F_x \frac{dx_1}{dt} = m_1 v_{x(1)} \frac{dv_{x(1)}}{dt}.$$

If the force F_x arises from the interaction of the mass m_1 with another body of mass m_2, the force acting on m_2 in the x-direction is $-F_x$ by the third law.

Applying the second law and multiplying by $\dfrac{dx_2}{dt}$, we get

$$-F_x \frac{dx_2}{dt} = m_2 v_{x(2)} \frac{dv_{x(2)}}{dt}.$$

We get four similar equations when we apply the same argument to F_y and F_z. Adding all these equations together and integrating with respect to t:

$$\int F_x d(x_2 - x_1) + \int F_y d(y_2 - y_1) + \int F_z d(z_2 - z_1)$$

$$+ \frac{m_1}{2}(v^2_{x(1)} + v^2_{y(1)} + v^2_{z(1)}) + \frac{m_2}{2}(v^2_{x(2)} + v^2_{y(2)} + v^2_{z(2)}) = \text{constant}$$

(the constant of integration).

The integral terms represent the work done by the force **F** in altering the distance apart of bodies m_1 and m_2, so that the change in the total *kinetic energy* $\frac{1}{2}m_1 v_{(1)}^2 + \frac{1}{2}m_2 v_{(2)}^2$ of m_1 and m_2 is equal to the work done by the forces. Further, the terms $\int F_x dx_2$, $\int F_y dy_2$, $\int F_z dz_2$, represent the work done by **F** on m_2, so that, if we chose to consider that m_1 was *outside* the assembly and reckoned the force **F** as an *external* one acting on m_2, we should conclude that the change in the kinetic energy of m_2 was equal to the external work done by the forces on m_2, and a similar result can be found for an assembly of many bodies subject to "external" forces.

The work done by forces, internal or external, acting on any system of bodies is usually called change of *potential* energy, and the principle is expressed by saying that the sum of kinetic and potential energies is constant. Since only *changes* of potential energy are defined, its zero is arbitrary.

(c) *The conservation of angular momentum*

The *angular momentum* of a set of massive particles about the x-axis, is defined as the sum

$$\sum_l m_l (y_l v_{z(l)} - z_l v_{y(l)})$$

where m_l is the mass of the lth particle, y_l its y-coordinate, $v_{y(l)}$ its y-component of velocity, etc. Clearly this sum depends upon the *position* of the x-axis as well as on its direction. Linear momentum only depends on the *directions* of the axes of coordinates, while energy is the same for any choice of axes. Consider now the effect of a force acting between the lth and mth particles, with components F_x, F_y, F_z acting on the lth particle. We have

$$F_y = m_l \frac{dv_{y(l)}}{dt} = m_l \frac{d^2y_l}{dt^2},$$

$$F_z = m_l \frac{dv_{z(l)}}{dt} = m_l \frac{d^2z_l}{dt^2}.$$

Then
$$y_l F_z - z_l F_y = m_l \left(y_l \frac{d^2z_l}{dt^2} - z_l \frac{d^2y_l}{dt^2} \right)$$
$$= m_l \frac{d}{dt}\left(y_l \frac{dz_l}{dt} - z_l \frac{dy_l}{dt} \right) \ldots\ldots\ldots\ldots \quad (1)$$

In the same way, consider the effect of the forces $-F_z$, $-F_y$, which, by the third law, must act on the mth particle. We have

$$-y_m F_z + z_m F_y = m_m \frac{d}{dt}\left(y_m \frac{dz_m}{dt} - z_m \frac{dy_m}{dt} \right).$$

So that the contribution to the rate of change of $\sum_l m_l (y_l v_{z(l)} - z_l v_{y(l)})$ due to the forces F_y, F_z is

$$(y_l - y_m) F_z - (z_l - z_m) F_y,$$

which vanishes if the force between particles l and m acts along the line joining them, so that, *if a system of particles is subject only to mutual central forces which obey Newton's third law, its total angular momentum remains constant.*

Equation (1) also holds if F_y and F_z are external forces acting on the lth particle, so that, by summing over all particles, we have

$$\frac{dM_x}{dt} = \frac{d}{dt}\left[\sum_l m_l \left(y_l \frac{dz_l}{dt} - z_l \frac{dy_l}{dt} \right) \right] = \sum_l (y_l F_z - z_l F_y)$$

that is, *the rate of change of angular momentum referred to a fixed axis is equal to the total moment of the external forces about that axis.*

It may be asked what happens to these three laws if the forces between particles do not obey Newton's third law, or are not central. Such forces certainly exist in nature, *e.g.* the magnetic forces between moving charged bodies. It will be noticed that, to prove the laws of conservation of energy and of linear momentum, we had to use only the condition that the mutual forces between two particles were equal and opposite, whereas to prove the conservation of angular momentum, we needed also the condition that the forces were central. Forces lacking one or both of these properties are called *non-conservative*, because the conservation laws need not hold for such forces.

We may be able to prove one or more of the conservation laws for *particular types* of non-conservative forces. As an example, the conservation of energy still holds even if magnetic interactions are present, because magnetic forces on a charge are always perpendicular to the direction in which the charge is moving, and therefore can neither do work on the charge nor alter the *magnitude* of its velocity, but can only change its *direction*. Such forces therefore do not affect either the potential or kinetic energies.

(*d*) *The principle of virtual work*

A special case of the law of conservation of energy can be used to deduce conditions of static equilibrium and is applicable to very complicated mechanical systems. As a simple case illustrating the method, consider a particle in equilibrium under the influence of a number of co-planar strings. We already know how to determine the condition of equilibrium by means of the parallelogram of forces (Fig. 1), or, what amounts to the same thing, by equating the sum of the components of force in the x-direction to zero, and then doing the same thing for some other direction. Now we may imagine instead the particle to be displaced slightly from its equilibrium position, and we also suppose that the tension in the strings are kept constant. This could be arranged if, for example, the strings ran over pulleys and were connected to weights. Then we express our condition of equilibrium by saying that, for any small displacement of the particle, *the total work done by the forces acting on it vanishes to the first order of small quantities*. For, if we suppose the particle displaced a distance dx along the x-axis, the corresponding shortening of a string inclined to the x-axis at an angle θ is $dx \cos \theta +$ terms of order $(dx)^2$ and the work associated with this change of length of the string is $T \cos \theta dx$, where T is the tension in the string. The total work associated with this displacement is then

$$dx \sum_l T_l \cos \theta_l + \text{terms of order } (dx)^2$$

where T_l is the tension in the lth string, and θ_l its inclination to the x-axis. The condition that the coefficient of dx is to vanish is the same as the condition that the total force along the x-axis is zero.

Expressed in another way, the principle states that if a system is in static equilibrium, its potential energy is a maximum or minimum, so that the work associated with any small displacement dx compatible with the geometry of the system is, at most, of order $(dx)^2$. We can deduce the conditions of equilibrium of the system by assuming different types of displacement. The principle is sometimes more convenient than the method of composition of forces when we are concerned with complicated systems or continuous media.

We can give a formal proof that, for a continuous medium, the principle gives results equivalent to those of ordinary methods of statics. Evidently *any* deformation of a continuous medium could, in principle, be realised by moving the particles to their new positions one at a time, keeping the others fixed during each step. Suppose that a typical particle is moved through a distance dr, its changes of coordinates being dx, dy, dz. Then, if $\mathbf{p}_k$ is one of the forces acting on this particle, the work done by this force during the supposed displacement is

$$p_{k(x)}\,dx + p_{k(y)}\,dy + p_{k(z)}\,dz$$

where $p_{k(x)}$ is the x-component of $\mathbf{p}_k$, etc.

Now the principle of virtual work states that this quantity, summed over all the forces acting on the typical particle, is zero for any displacement, that is

$$dx \sum_k p_{k(x)} + dy \sum_k p_{k(y)} + dz \sum_k p_{k(z)} = 0$$

for any values of dx, dy, dz, which is only possible if $\sum_k p_{k(x)}$, $\sum_k p_{k(y)}$, $\sum_k p_{k(z)}$, all vanish for every particle. But these are just the ordinary conditions of static equilibrium, that the resultant of all the forces acting at any point is zero. We have made the implicit assumption that the work done is the same in whatever *order* we displace the particles, and this can be shown to be equivalent to assuming that the forces between two particles act only along the line joining them, *i.e.* that the forces are central and therefore conservative (p. 12).

The principle is of value for a complicated system because we can choose our virtual displacement in such a way that we avoid having to calculate some of the forces. Thus, consider the framework shown in Fig. 2, consisting of six freely-jointed rods, each of weight mg, AF being fixed and the others suspended in a vertical plane. Now it is quite possible to calculate the tension in the string GH by the ordinary methods of resolving forces and taking moments, but this involves considering the forces *along* the rods and the reactions between the rods at the corners. We can deal with the problem much more neatly by the principle of virtual work, by supposing the member DC pushed vertically upwards, the angles FAB and BCD becoming $120° + d\theta$. With such a displacement the reactions at the hinges do no work and need not be considered explicitly, and a similar remark applies to the thrusts or tensions in the rods themselves. (These forces do no work because, in our assumed displacement, the rods remain in contact at the joints, which means that the reactions do no work, and of invariable length, which means that the thrusts or tensions also do no work.) Regarding the tension in the string, we make use of the device already referred to, and suppose G replaced by a smooth pulley, carrying a mass of weight equal to T at the other end. Then if, in Fig. 2, BH' represents AB produced, H' will rise as much as H, and L' as much as L. Then, as a result of the rotation through $d\theta$ of AB, K is lifted by a distance $-d\,(a\cos\theta) = a\sin\theta d\theta = \frac{\sqrt{3}a d\theta}{2}$, if the hexagon is regular initially. L' and H' are lifted respectively by 3 and 4 times this distance, and the same is true of L and H. Therefore we have:

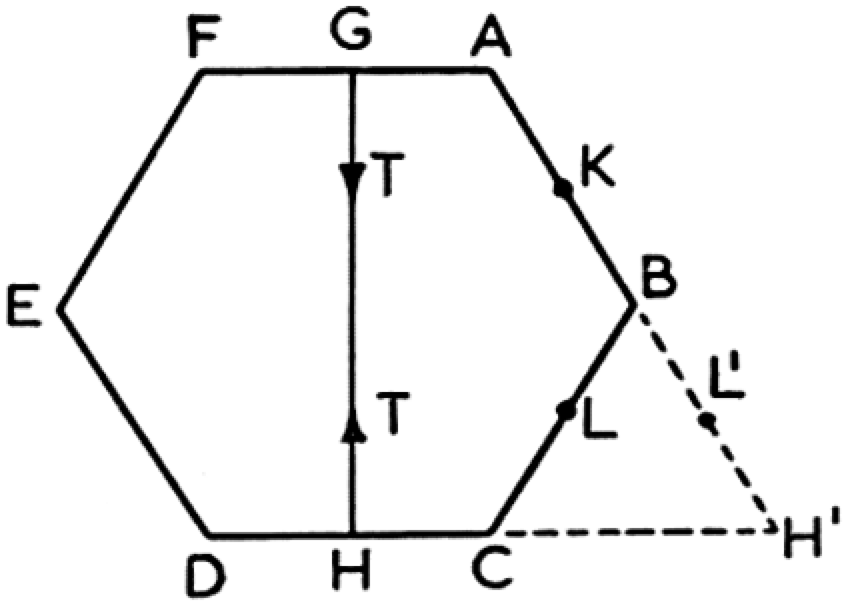


FIG. 2.

Work done in lifting centres of gravity of AB and $EF = 2 . \frac{\sqrt{3}a d\theta}{2} . mg$

,, ,, ,, ,, BC and $DE = 2.3.\frac{\sqrt{3}a d\theta}{2} . mg$

,, ,, centre of gravity of $CD = 4 . \frac{\sqrt{3}a d\theta}{2} . mg$

This is to be equal to the work done by the tension in the string, which, since H has risen a distance $\frac{4\sqrt{3}a d\theta}{2}$ is $T . \frac{4\sqrt{3}a d\theta}{2}$. Thus we have

$$4T = mg\,(2 + 6 + 4) \quad \text{or} \quad T = 3mg.$$

Types of force that can evidently be eliminated from explicit consideration by a skilful choice of virtual displacement are, as we have seen, thrusts or tensions in rods of invariable length, tensions in inextensible strings (unless we wish to calculate the tensions explicitly, when the artifice used above can be applied), and reactions between a body and a smooth plane. The latter does no work if we restrict ourselves to displacements *along* the plane. (If it is a *rough* plane, the work done against friction must be added into the equation of virtual work.)

As a second example of the great power of this method, we shall consider the common letter balance (Fig. 3). This is in equilibrium whenever the sum of the mass M_1, placed in one pan, and the mass m_1 of the pan itself and its upright member, is equal to the corresponding sum for the other pan, regardless of the positions in the pans. This would be a distinctly troublesome result to prove by straightforward resolution and taking moments, but follows at once from the principle of virtual work. We need only consider a small rotation of ABC about B. The centres of gravity of ABC and DEF remain stationary, the work done on AD and its scale pan is m_1gdh, that on M_1 is M_1gdh, while that done by the weight of CF and its scale pan is m_2gdh, and that done by the weight of M_2 is M_2gdh. Simple geometrical considerations show us that the upward and downward displacements of AD, CF, M_1, and M_2 are all dh, so that the condition of equilibrium is

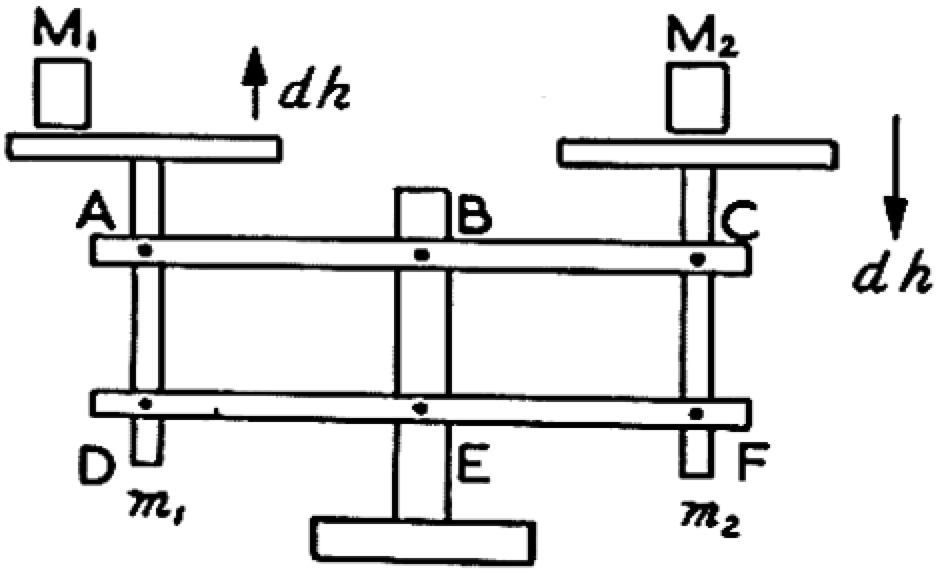


Fig. 3.

$$M_1 + m_1 = M_2 + m_2,$$

a result that is by no means obvious or elementary. We must, of course, presume that $AB = BC = DE = EF$, and also that the centres of gravity of AC and DF are at B and E. If these conditions were not satisfied, we could still determine the laws of such a "false" balance by means of the principle of virtual work very much more readily than by other means.

4. Laws of one-dimensional motion

Since acceleration is proportional to force, it is possible to calculate the motion of a body of fixed mass if the force is known as a function of time. We have

$$m\frac{d^2x}{dt^2}=f(t).$$

Integrating twice, we get

$$x=\frac{1}{m}\int[\int f(t)\ dt]\,dt+At+B$$

where A and B are constants determined, *e.g.* if we know x and dx/dt at a given time. It usually happens that the force is not given explicitly as a function of the time t, for it may also depend on the position x, or the velocity dx/dt. In this case, we have to solve a *differential equation* in order to determine the motion, and we now consider some typical cases.

(*a*) *Motion in a resisting medium*

If the resistance is proportional to the velocity, the equation of motion is

$$m\frac{d^2x}{dt^2}+k\frac{dx}{dt}=f(t).$$

Integrating once we obtain

$$\frac{dx}{dt}+\frac{k}{m}x=\frac{1}{m}\int f(t)\,dt+A.$$

Multiply through by $e^{kt/m}$

$$e^{kt/m}\frac{dx}{dt}+\frac{k}{m}e^{kt/m}x=\frac{e^{kt/m}}{m}\int f(t)\,dt+Ae^{kt/m}.$$

The left-hand side can now be integrated, and we obtain:

$$e^{kt/m}x=\frac{1}{m}\int e^{kt/m}dt\int f(t)\,dt+\frac{mA}{k}e^{kt/m}+B\ \ldots\ldots\ldots\ (2)$$

so that the motion can be calculated for any form of the function $f(t)$, A and B being determined by the initial conditions, *e.g.* if we know the values of x and dx/dt at a given time.

An important special case occurs when $f(t)$ is a constant, the force being due, *e.g.* to gravity. In the latter case, we may put $f(t)$ equal to $-mg$, where g is the acceleration of gravity (see Chapter III) and equation (2) now becomes, after performing the integrations,

$$x=-\frac{gt}{k}+\frac{mg}{k^2}+\frac{mA}{k}+Be^{-kt/m}.$$

Without loss of generality, we may suppose the origin so chosen that $x = 0$ at $t = 0$, which gives us $B + \frac{mA}{k} + \frac{mg}{k^2} = 0$. From equation (2) putting $t = 0$, we find that A must be equal to v_0, the initial velocity, so that we obtain as the complete solution

$$x = -\frac{gt}{k} + \left(\frac{mv_0}{k} + \frac{mg}{k^2}\right)(1 - e^{-kt/m}) \ldots\ldots\ldots\ldots \quad (3)$$

Differentiating with respect to time,

$$\frac{dx}{dt} = -\frac{g}{k} + \left(v_0 + \frac{g}{k}\right)e^{-kt/m} \ldots\ldots\ldots\ldots \quad (4)$$

As t increases, the exponential terms gradually become smaller, and equation (4) shows that for any initial velocity v_0 (upwards or downwards), the body finally attains a downward *terminal* velocity g/k.

If a body moves under a constant force, such as gravity, but is subject to a resistance proportional to the square of the velocity, its motion can be calculated in the following way. The equation of motion is

$$m\frac{d^2x}{dt^2} + k\left(\frac{dx}{dt}\right)^2 = g$$

which may also be written in terms of the velocity v

$$\frac{dv}{dt} = \frac{g - kv^2}{m} \ldots\ldots\ldots\ldots \quad (5)$$

or

$$\int dt = m\int\frac{dv}{g - kv^2}$$

by separating the variables and integrating both sides of equation (5).

$$C + t = \frac{m}{\sqrt{gk}}\tanh^{-1}\left(v\sqrt{\frac{k}{g}}\right)$$

or

$$v = \sqrt{\frac{g}{k}}\tanh\left[\frac{\sqrt{gk}}{m}(t + C)\right] \ldots\ldots\ldots\ldots \quad (6)$$

where C may be determined if v is known at some time.

This may again be integrated to give:

$$x = \frac{m}{k}\log\cosh\left[\frac{\sqrt{gk}}{m}(t + C)\right] + D$$

where D is a constant of integration which may be determined if x is known at some time. Equation (6) shows that, as t becomes large and

$\tanh\left[\frac{\sqrt{gh}}{m}(t+C)\right]$ approaches unity, v approaches a terminal value $\sqrt{g/k}$, when the force of gravity becomes equal to the resistance. Again the terminal velocity is independent of the initial value of v.

(*b*) *Motion under elastic constraint. Simple harmonic motion*

It may happen that the force acting on a body is independent of its velocity, but is proportional to its distance from a certain fixed point. Such motion is extremely common, occurring when *any* system in stable equilibrium is slightly disturbed. If we take the fixed point as $x = 0$, the equation of motion is:

$$m\frac{d^2x}{dt^2} = -\lambda x$$

where λ is the ratio of force to displacement. The negative sign means the elastic constraint tends to diminish x. If we multiply both sides of this equation by dx/dt it may be integrated to give

$$\tfrac{1}{2}m\left(\frac{dx}{dt}\right)^2 = -\tfrac{1}{2}\lambda x^2 + E \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (7)$$

where E is a constant.

Let x_0 be the value of x for which the velocity dx/dt is zero. Then $E = \frac{1}{2}\lambda x_0^2$ and equation (7) becomes

$$\frac{dx}{dt} = \sqrt{\frac{\lambda}{m}}(x_0^2 - x^2)^{\frac{1}{2}} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (8)$$

which may be written

$$dt = \sqrt{\frac{m}{\lambda}}\frac{dx}{(x_0^2 - x^2)^{\frac{1}{2}}}.$$

Integrating both sides

$$t + F = \sqrt{\frac{m}{\lambda}}\sin^{-1}\left[\frac{x}{x_0}\right]$$

$$\text{or } x = x_0\sin\left[\sqrt{\frac{\lambda}{m}}(t+F)\right] \text{ and } \frac{dx}{dt} = x_0\sqrt{\frac{\lambda}{m}}\cos\left[\sqrt{\frac{\lambda}{m}}(t+F)\right] \quad (9)$$

using equation (8).

Since $\qquad \sin\theta = \sin(2\pi + \theta)$

and $\qquad \cos\theta = \cos(2\pi + \theta)$

it follows that x and dx/dt return to the same values whenever $t\sqrt{\lambda/m}$ increases by 2π, that is *the motion is periodic with a periodic time* $T = 2\pi\sqrt{m/\lambda}$. This result holds whatever may be the values of x_0 and F, that is, *the periodic time is independent of the amplitude of the* *motion and of the initial values of x and dx/dt.* This is the characteristic feature of simple harmonic motion.

The more complicated cases of the motion of a body under a force proportional to its displacement, when it is subject also to a resistance depending on its velocity and to an external periodic force, are treated in the Appendix (p. 285).

5. Motion in two dimensions

(a) *Trajectories*

We consider now the motion of a body, subject to no force but gravity, which is given an initial velocity in some direction other than the vertical. We take the y-axis as vertical and suppose that gravity acts downwards, taken to be the negative direction. We may then suppose that the velocity of the body at each instant is resolved into components in the x and y directions, according to the parallelogram law. If v_0 is the initial velocity, inclined at an angle a to the horizontal, the horizontal component of velocity is $v_0 \cos a$, which remains constant, since the acceleration of gravity is vertical. We then have

$$x = v_0 t \cos a \quad \text{if} \quad x = 0 \text{ at } t = 0 \quad \ldots\ldots\ldots\ldots \quad (10)$$

The equation of motion in the vertical direction is:

$$\frac{d^2y}{dt^2} = -g.$$

Integrating, $\dfrac{dy}{dt} = -gt + v_0 \sin a \left(\text{since } \dfrac{dy}{dt} = v_0 \sin a \text{ at } t = 0\right).$

Integrating again, $y = -\tfrac{1}{2}gt^2 + v_0 t \sin a$ (if $y = 0$ at $t = 0$).

Substituting for t from equation (10) we get for the equation of the trajectory:

$$y = x \tan a - \tfrac{1}{2}\frac{gx^2 \sec^2 a}{v_0^2} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (11)$$

which is a parabola. This parabola cuts the x-axis at the origin and at the point

$$x = \frac{2v_0^2}{g} \sin a \cos a = \frac{v_0^2}{g} \sin 2a.$$

The distance between these two points (the range of a projectile on a flat plane) reaches a maximum value v_0^2/g when $a = \pi/4$.

If we suppose that the body is moving in a medium that exerts a resisting force kv proportional to the velocity v, and in the same direction as the velocity, we may resolve the force kv into components $kv \cos \theta$, $kv \sin \theta$ parallel to the x and y axes respectively, where θ is the angle at which the velocity is inclined to the x-axis. The equation of motion in the horizontal direction is now:

$$m\frac{d}{dt}(v\cos\theta) = -k(v\cos\theta) \quad \ldots\ldots\ldots\ldots \quad (12)$$

while

$$m\frac{d}{dt}(v\sin\theta) = -k(v\sin\theta) - mg \quad \ldots\ldots\ldots\ldots \quad (13)$$

is the equation of motion in the vertical direction. Now $v\cos\theta$, $v\sin\theta$ are the horizontal and vertical components of the velocity, and may therefore be replaced by dx/dt, dy/dt respectively. These equations may be integrated exactly as in Art. 4 (a), with $f(t)$ put equal to zero for the x-equation (12) and equal to $-mg$ for the y-equation (13). We thus obtain

$$x = \frac{mA}{k} + Be^{-kt/m} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (14)$$

$$y = -\frac{gt}{k} + \frac{mg}{k^2} + \frac{mC}{k} + De^{-kt/m} \quad \ldots\ldots\ldots\ldots \quad (15)$$

where A, B, C, and D are constants that have to be determined from the initial conditions (usually the two initial coordinates and two components of velocity of the body). If we suppose that the body starts from the origin at $t = 0$,

then $$B = -\frac{mA}{k}, \quad D = -\frac{mC}{k} - \frac{mg}{k^2}.$$

The two remaining constants may be determined if we know the magnitude and direction of the initial velocity. Suppose it has magnitude v_0, inclined at a to the horizon. Differentiating equations (14) and (15) with respect to t

$$\frac{dx}{dt} = -\frac{kB}{m}e^{-kt/m}$$

so that at $t = 0$, $$v_0\cos a = \frac{dx}{dt} = -\frac{kB}{m}$$

$$\frac{dy}{dt} = \frac{-g}{k} - \frac{kD}{m}e^{-kt/m}$$

so that at $t = 0$, $$v_0\sin a = \frac{dy}{dt} = \frac{-g}{k} - \frac{kD}{m}.$$

In inserting the values thus obtained for A, B, C, and D, we obtain finally:

$$x = \frac{mv_0}{k}\cos a\,(1 - e^{-kt/m}) \quad \ldots\ldots\ldots\ldots\ldots \quad (16)$$

$$y = \frac{mv_0}{k}\sin a\,(1 - e^{-kt/m}) - \frac{gt}{k} + \frac{mg}{k^2}(1 - e^{-kt/m}) \quad \ldots\ldots \quad (17)$$

and the equation of the trajectory can be obtained by eliminating t between these two equations. Since the resistance actually encountered by a projectile is much more nearly proportional to v^2, detailed calculations from these equations (16), (17) are not of great practical use. If the resistance is supposed proportional to v^2, the equations of motion in the two-dimensional case can be integrated only numerically.

(*b*) *Motion in a central field of force. Orbits*

If a body moves under the influence of a force which is always directed towards one point, it is said to describe a *central orbit.* Usually, such a force depends only on the distance of the body from the fixed point, and not on the direction of the line joining it to the fixed point. It is convenient to describe the position of the body by means of polar coordinates referred to the fixed point as origin. The direction of the initial velocity of the body, together with the fixed point, determine a plane, and since the force on the body lies along the line joining it to the fixed point, that is, in this plane, there can be no force perpendicular to the plane, and the motion is always in that plane, so that the problem is a two-dimensional one.

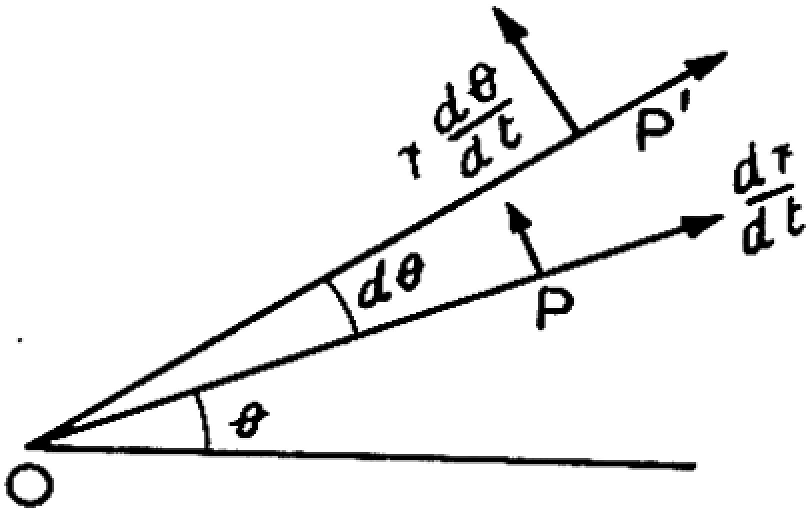


FIG. 4.

At any instant we can resolve the velocity of the body into a component $\frac{dr}{dt}$ along r and a component $r\frac{d\theta}{dt}$ perpendicular to r (Fig. 4). We now need expressions for the *accelerations* in these two directions. In a short time dt, the angle θ will have increased by $d\theta$, but the acceleration along r is not simply d^2r/dt^2, because the *directions* of the velocities along and perpendicular to r have changed as well as their magnitudes. To calculate the radial acceleration we must resolve the new velocities after the time dt along the *original* direction of r, and then examine how much the velocity *in this direction* has changed from the original value dr/dt. Thus the change in radial velocity is

$$\left(\frac{dr}{dt}+dt\frac{d^2r}{dt^2}\right)\cos(d\theta)-\sin(d\theta)\left[r\frac{d\theta}{dt}+dt\frac{d}{dt}\left(r\frac{d\theta}{dt}\right)\right]-\frac{dr}{dt}.$$

Similarly, the change in transverse velocity in the direction perpendicular to the original direction of r is

$$\left[r\frac{d\theta}{dt}+dt\frac{d}{dt}\left(r\frac{d\theta}{dt}\right)\right]\cos(d\theta)+\left(\frac{dr}{dt}+dt\frac{d^2r}{dt^2}\right)\sin(d\theta)-r\frac{d\theta}{dt}.$$

For $d\theta$ very small, cos $(d\theta)$ may be replaced by unity, and sin $(d\theta)$ by $d\theta$. Thus, by evaluating these expressions, dividing by dt, and proceeding to the limit when $d\theta$ and dt become very small, we obtain:

$$\text{Radial acceleration} = \frac{d^2r}{dt^2} - r\left(\frac{d\theta}{dt}\right)^2 \quad \ldots\ldots\ldots\ldots\ldots \quad (18)$$

$$\text{Transverse acceleration} = \frac{d}{dt}\left(r\frac{d\theta}{dt}\right) + \frac{dr}{dt}\cdot\frac{d\theta}{dt}$$

$$= r\frac{d^2\theta}{dt^2} + 2\frac{dr}{dt}\cdot\frac{d\theta}{dt} \quad \ldots\ldots\ldots\ldots\ldots \quad (19)$$

The term $-r\left(\frac{d\theta}{dt}\right)^2$ is often called the *centrifugal acceleration*, while the term $2\frac{dr}{dt}\cdot\frac{d\theta}{dt}$ is called the *Coriolis acceleration* after its discoverer.

Applying Newton's second law, it is clear that, for a force $f(r)$ always in the direction of r and depending only on its length, the radial acceleration may be expressed as $f(r)/m$, where m is the mass of the body if an inward or attractive force is reckoned positive, while the transverse acceleration is always zero, because the force acting on the body is along r and therefore always perpendicular to the transverse velocity. The equations determining r and θ are therefore

$$\frac{d^2r}{dt^2} - r\left(\frac{d\theta}{dt}\right)^2 = -\frac{f(r)}{m} \quad \ldots\ldots\ldots\ldots\ldots \quad (20)$$

$$r\frac{d^2\theta}{dt^2} + 2\frac{dr}{dt}\cdot\frac{d\theta}{dt} = 0 \quad \ldots\ldots\ldots\ldots\ldots \quad (21)$$

Equation (21), after multiplication through by r, may be written

$$\frac{d}{dt}\left(r^2\frac{d\theta}{dt}\right) = 0 \quad \text{or} \quad r^2\frac{d\theta}{dt} = \text{constant} \quad \ldots\ldots\ldots \quad (22)$$

Thus, the quantity $r^2\frac{d\theta}{dt}$ is always equal to the value it had initially, and it is sometimes known as the *impact parameter* because it vanishes if the initial velocity of the body points directly towards the centre of force. The quantity $mr^2d\theta/dt$ is the *angular momentum* of the moving body about an axis through the central point perpendicular to the plane of the orbit, and equation (22) is equivalent to saying that *angular momentum is conserved during the motion.*

Another way of expressing equation (22) is by the statement that *the body sweeps out the area of the orbit at a uniform rate,* because an element of area of the orbit included between the lines corresponding to θ and $\theta + d\theta$ is $\frac{1}{2}r^2 d\theta$. This is the second of the laws of planetary motion, discovered empirically by **Kepler**. If k is the constant value of $r^2\dfrac{d\theta}{dt}$, then we have from equation (20)

$$\frac{d^2r}{dt^2} - \frac{k^2}{r^3} = -\frac{f(r)}{m}.$$

Multiplying through by dr/dt, this equation can be integrated

$$\frac{1}{2}\left(\frac{dr}{dt}\right)^2 + \frac{1}{2}\frac{k^2}{r^2} = \frac{1}{m}\int_r^{\infty} f(r)\,dr + A \qquad \text{(23)}$$

which might have been deduced directly from the equation of conservation of energy (p. 10) because dr/dt is the radial velocity of the body, while k/r is $r\dfrac{d\theta}{dt}$ or the transverse velocity and $\displaystyle\int_{r_1}^{r_2} f(r)$ is the work done by the attractive force $f(r)$ as r changes from r_1 to r_2. We may now use equation (22) to change the independent variable from t to θ, thus obtaining an equation for the shape of the orbit. We have

$$\frac{dr}{dt} = \frac{dr}{d\theta}\cdot\frac{d\theta}{dt} = \frac{k}{r^2}\cdot\frac{dr}{d\theta}$$

so that equation (24) becomes

$$\frac{1}{r^4}\left(\frac{dr}{d\theta}\right)^2 + \frac{1}{r^2} = \frac{2}{mk^2}\int_r^{\infty} f(r)\,dr + \frac{2A}{k^2} \qquad \text{(24)}$$

This equation is true whatever the form of $f(r)$, provided only that the force is always directed towards the origin. To make any further progress we must know the form of $f(r)$. Two special cases are of great importance.

Case I.—$f(r) = K/r^2$ (the inverse-square law).

Equation (25) may now be written

$$\frac{1}{r^4}\left(\frac{dr}{d\theta}\right)^2 + \frac{1}{r^2} = -\frac{2K}{mk^2 r} + \frac{2A}{k^2} \qquad \text{(25)}$$

This equation simplifies somewhat if we put $u = 1/r$.

It may be written

$$\left(\frac{du}{d\theta}\right)^2 + u^2 = -\frac{2K}{mk^2}u + \frac{2A}{k^2},$$

or

$$\frac{du}{d\theta} = \left(\frac{2A}{k^2} - \frac{2K}{mk^2}u - u^2\right)^{\frac{1}{2}},$$

or
$$d\theta = \frac{du}{\left[\dfrac{2A}{k^2} + \left(\dfrac{K}{mk^2}\right)^2 - \left(u + \dfrac{K}{mk^2}\right)^2\right]^{\frac{1}{2}}}.$$

Integrating
$$\theta + \beta = \cos^{-1}\left[\frac{u + \dfrac{K}{mk^2}}{\left(\dfrac{2A}{k^2} + \dfrac{K^2}{m^2k^4}\right)^{\frac{1}{2}}}\right],$$

or
$$u = \left(\frac{2A}{k^2} + \frac{K^2}{m^2k^4}\right)^{\frac{1}{2}} \cos(\theta + \beta) - \frac{K}{mk^2} \quad \ldots\ldots\ldots \quad (26)$$

where β is another constant of integration.

Putting $u = 1/r$, we find that the orbit is a conic with a focus at the centre of attraction, with major axis lying along the line $\theta = -\beta$, semi-latus rectum mk^2/K, and eccentricity
$$\left(1 + \frac{2m^2Ak^2}{K^2}\right)$$
because the polar equation of a conic of eccentricity e and semi-latus rectum l, with a focus at the origin and a major axis lying along the initial line, is
$$\frac{l}{r} = 1 + e\cos\theta.$$

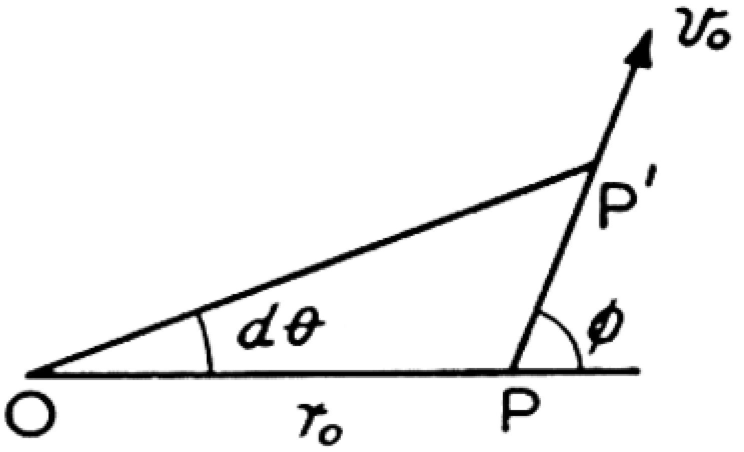


Fig. 5.

Thus the shape of the orbit is completely determined provided that the constants k, A, and B are known. These can be determined provided that we know the initial position of the body, and the magnitude and direction of the initial velocity. Without loss of generality, we may suppose the initial line for the polar coordinates to pass through P, the initial position of the body. Let the direction of its velocity then be as in Fig. 5, where v_0 is the initial velocity, inclined at an angle ϕ to the initial line. After a short time dt, the body has moved from P to P' and its θ coordinate is then the angle $POP' = d\theta$. For $d\theta$ small we have
$$\frac{PP'}{\sin(d\theta)} = \frac{OP'}{\sin\phi} \quad \text{or} \quad \left(\frac{d\theta}{dt}\right)_{t=0} = \frac{v_0 \sin\phi}{r_0}$$

So that
$$k = r_0^2\left(\frac{d\theta}{dt}\right)_{t=0} = r_0 v_0 \sin\phi.$$

Now $r_0 \sin\phi$ is the perpendicular distance between O and the initial direction of v, so that k vanishes only if the velocity is directly towards or away from O. To determine A we return to equation (23) which becomes

$$A = \frac{1}{2}\left(\frac{dr}{dt}\right)^2 + \frac{1}{2}\frac{k^2}{r^2} - \frac{l}{m}\frac{K}{r}$$

for the particular case of the inverse-square law. As we have already pointed out

$$\left(\frac{dr}{dt}\right)^2 + \frac{k^2}{r^2} \quad \text{is equal to} \quad \left(\frac{dr}{dt}\right)^2 + r^2\left(\frac{d\theta}{dt}\right)^2,$$

that is to the square of the total velocity. We have at the initial instant

$$A = \frac{1}{2}v_0^2 - \frac{K}{mr_0},$$

so that mA is the difference between $\frac{1}{2}mv_0^2$, the initial kinetic energy of the body, and K/r_0, which represents the work necessary to remove the body from its initial position to infinity against the attractive force. We should expect quite a different type of orbit according as the body has, or has not, enough kinetic energy to enable it to escape eventually from the influence of the attractive force, that is according as A is positive or negative. We may distinguish the following cases.

If $k = 0$, the orbit is a straight line, and the particle either oscillates about the centre of attraction (the so-called "pendulum orbit") or travels away to infinity, when A is negative or positive respectively.

If $k > 0$ but $A < 0$, we have $e < 1$, and the orbit is an ellipse.

If $k > 0$ but $A = 0$, we have $e = 1$, and the orbit is a parabola.

If $k > 0$ but $A > 0$, we have $e > 1$, and the orbit is a hyperbola.

If $k > 0$ and $A = -\dfrac{K^2}{2m^2k^2}$, $e = 0$, and the orbit is a circle.

Since $A = \frac{1}{2}v_0^2 - K/mr_0$ and $k = r_0v_0 \sin\phi$, we have

$$\tfrac{1}{2}v_0^2 - \frac{K}{mr_0} + \frac{K^2}{2m^2r_0^2v_0^2\sin^2\phi} = 0$$

if the orbit is to be a circle, which is compatible with a real value of v_0^2 only if $\sin\phi$ is unity, *i.e.* if the line PP' (Fig. 5) is perpendicular to OP.

It remains only to determine β, the inclination of OP to the major axis. From Fig. 5, the initial value of $dr/d\theta$ is

$$\frac{OP' - OP}{\angle POP'} = \frac{PP'}{d\theta}\cos\phi = \frac{r_0}{\sin\phi}\cos\phi = r_0\cot\phi.$$

From equation (26) by differentiation we have

$$-\frac{1}{r^2}\frac{dr}{d\theta} = -\left(\frac{2A}{k^2} + \frac{K^2}{m^2k^4}\right)^{\frac{1}{2}}\sin(\theta + \beta).$$

Initially $\theta = 0$ and $r = r_0$, so that β may be determined if the other constants are known.

If the orbit is elliptic it will repeat itself as θ changes by 2π, and we need to determine the corresponding periodic time. Since the area of an ellipse is πab, where a and b are the major and minor semi-axes and this area is described at the rate $k/2$ (p. 23), the periodic time T is $2\pi ab/k$. Since the semi-latus rectum l is b^2/a we have

$$T^2 = \frac{4\pi^2 a^2 b^2}{k^2} = \frac{4\pi^2 a^3 l}{k^2} = \frac{4\pi^2 a^3 m}{K}$$

since l has been found to be equal to mk^2/K (p. 24). *Thus the square of the periodic time of an elliptic orbit* (*for a fixed* m *and* K) *is proportional to the cube of its semi-major axis* (Kepler's third law). Since the length of a hyperbolic or parabolic orbit is infinite, such orbits have no finite period.

Case II.—Elastic constraint. $f(r) = \lambda r$.

This case can also be handled by the method used for Case I, but the results are cumbersome, and it is better to use Cartesian coordinates, for, if the force is proportional to the radial coordinate, it can be resolved into components λx, λy along the x and y axes, and the equations of motion are

$$m\frac{d^2x}{dt^2} + \lambda x = 0$$

$$m\frac{d^2y}{dt^2} + \lambda y = 0$$

which equations have already been solved (p. 18), the solutions being

$$x = x_0 \cos\left[\sqrt{\frac{\lambda}{m}}\,(t + A)\right]$$

$$y = y_0 \cos\left[\sqrt{\frac{\lambda}{m}}\,(t + B)\right]$$

where the constants x_0, y_0, A, B can be determined from the initial values of x, y, dx/dt, and dy/dt, that is, the initial position and velocity of the body. If A and B are equal, the ratio of x and y is always x_0/y_0, so that the orbit is a straight line. If $\sqrt{\lambda/m}\,(A - B)$ is an odd multiple of $\pi/2$, then

$$x = x_0 \cos\left[\sqrt{\frac{\lambda}{m}}\,(t + A)\right], \quad y = \pm\, y_0 \sin\left[\sqrt{\frac{\lambda}{m}}(t + A)\right]$$

and the orbit is an ellipse with major and minor semi-axes x_0 and y_0 respectively. If A and B differ by some other value, it can be shown (by rotating the coordinate axes) that the orbit is still an ellipse, but with its major axis inclined to the axes of coordinates. The period of the orbit is always $2\pi\sqrt{m/\lambda}$, independent of the initial conditions and the amplitudes x_0 and y_0.

Only particular types of variation of the attractive force with distance lead to *periodic* orbits and, in general, the solution of the differential equation of the orbit, although it may correspond to an orbit that constantly encircles the origin, implies some *progressive* change in the orbit as well. The best known example of this is Einstein's correction to Newton's inverse-square law of gravitational attraction (p. 69), which leads to a steady change in the direction of the major axis of the orbit with time, usually referred to as the *advance of perihelion.*

6. Motion of rigid bodies about a fixed axis. Pendulums

(a) *The simple pendulum*

We consider a particle of mass m attached to a horizontal pivot by a light rigid rod of length l, displaced from the vertical by an angle a, and allowed to swing under gravity [Fig. 6(a)].

If its inclination to the vertical at any instant is θ, the particle is moving instantaneously at right angles to the rod, and the force in this direction is $mg \sin \theta$, due to its weight, while its velocity in this direction is $\frac{d\theta}{dt}$. Its equation of motion is therefore

$$ml\frac{d^2\theta}{dt^2} + mg \sin \theta = 0 \quad \ldots\ldots\ldots\ldots \quad (27)$$

If a is very small, $\sin \theta$ may be replaced by θ, and equation (27) then becomes of the same form as the equation of simple harmonic motion already met, and the period T is given by

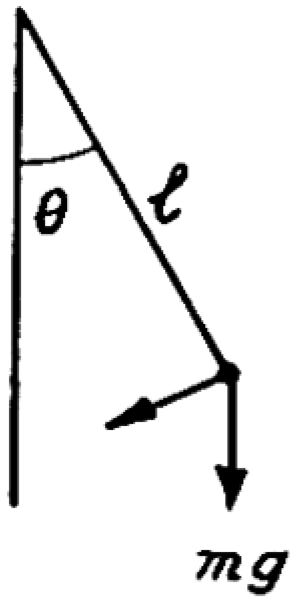


FIG. 6(a).

$$T = 2\pi\sqrt{\frac{l}{g}}.$$

We now ask if this period is affected if a is finite. Multiplying equation through (27) by $d\theta/dt$ and integrating we get

$$\tfrac{1}{2}ml\left(\frac{d\theta}{dt}\right)^2 - mg \cos \theta = \text{constant}$$

which might have been obtained directly from the principle of conservation of energy. If α is the maximum value of θ, $d\theta/dt$ is zero, so that the equation is

$$\tfrac{1}{2}ml\left(\frac{d\theta}{dt}\right)^2 = mg\,(\cos\theta - \cos\alpha);$$

$$\therefore \sqrt{\frac{l}{2g}}\,\frac{d\theta}{(\cos\theta - \cos\alpha)^{\frac{1}{2}}} = dt.$$

A complete swing is obtained when θ changes from $-\alpha$ to α and back to $-\alpha$; that is

$$T = \sqrt{\frac{2l}{g}}\int_{-\alpha}^{+\alpha}\frac{d\theta}{(\cos\theta - \cos\alpha)^{\frac{1}{2}}}.$$

This integral can be evaluated in terms of elementary functions only if $\cos\alpha \simeq 1$, which corresponds to the case of very small amplitude already treated, and the case $\cos\alpha = -1$, which corresponds to the pendulum being released from the position vertically above its support. In this case we have

$$T = \sqrt{\frac{2l}{g}}\int_{-\pi}^{+\pi}\frac{d\theta}{\sqrt{2}\cos\theta/2}$$

$$= 4\sqrt{\frac{l}{g}}\int_{0}^{\pi/2}\sec\phi d\phi = \infty$$

so that we may expect the period to lengthen as α increases. An approximate expression for T when α is small, but not zero, can be obtained as follows:

$$T = \sqrt{\frac{2l}{g}}\int_{-\alpha}^{+\alpha}\frac{d\theta}{(\cos\theta - \cos\alpha)^{\frac{1}{2}}}$$

$$= \sqrt{\frac{l}{g}}\int_{-\alpha}^{+\alpha}\frac{d\theta}{(\sin^2\alpha/2 - \sin^2\theta/2)^{\frac{1}{2}}}.$$

Put $\sin\frac{\theta}{2} = \sin\frac{\alpha}{2}\sin\phi$. Then $d\theta\cos\frac{\theta}{2} = 2d\phi\sin\frac{\alpha}{2}\cos\phi$

$$T = 2\sqrt{\frac{l}{g}}\int_{-\alpha/2}^{+\pi/2}\frac{d\phi}{\cos\theta/2} = 2\sqrt{\frac{l}{g}}\int_{-\pi/2}^{+\pi/2}\frac{d\phi}{(1 - \sin^2\alpha/2\,\sin^2\phi)^{\frac{1}{2}}}$$

$$= 4\sqrt{\frac{l}{g}}\int_{0}^{\pi/2} d\phi\,(1 + \tfrac{1}{2}\sin^2\phi\sin^2\alpha/2 + \tfrac{3}{8}\sin^4\phi\sin^4\alpha/2 + \ldots)$$

$$= 2\pi\sqrt{\frac{l}{g}}(1 + \alpha^2/16 + \ldots).$$

This gives the form of the correction when α is small. For other values of α the integral has been tabulated (it is known as a complete elliptic integral). See, for example, *Jahnke* and *Emde*, Functionentafeln.

(*b*) *The compound pendulum*

We consider now a rigid body pivoted about a horizontal axis [Fig. 6(*b*)].

Suppose that an element of the body of mass dm is at a distance r_m from the axis, the line from dm perpendicular to the axis, making an angle ϕ_m with the vertical. We may apply the principle of conservation of angular momentum to the whole body. There is no change in angular momentum only if the total moment of the weights of all the elementary masses dm is zero. The condition for this is

$$\Sigma dm\, r_m \sin\phi_m = 0 \quad \dots\dots\dots\dots\dots\dots \quad (28)$$

which means that, in equilibrium the body settles down in such a position that the average value of the coordinate $r_m \sin\phi_m$, which represents the horizontal distance of dm from the pivot, is zero. This is another way of saying that the *centre of gravity* lies vertically below the pivot.

Suppose that the body is rotated through a small angle θ about the axis. The velocity of the mass dm is then $r_m d\theta/dt$ at right angles to the direction of r_m, and the angular momentum of the body is then $\Sigma dm\, r_m{}^2 d\theta/dt$. The rate of decrease of this has to be equated to the moment of the weights of all the masses dm, and this is obtained by increasing all the angles ϕ_m in expression (28) by θ. This gives

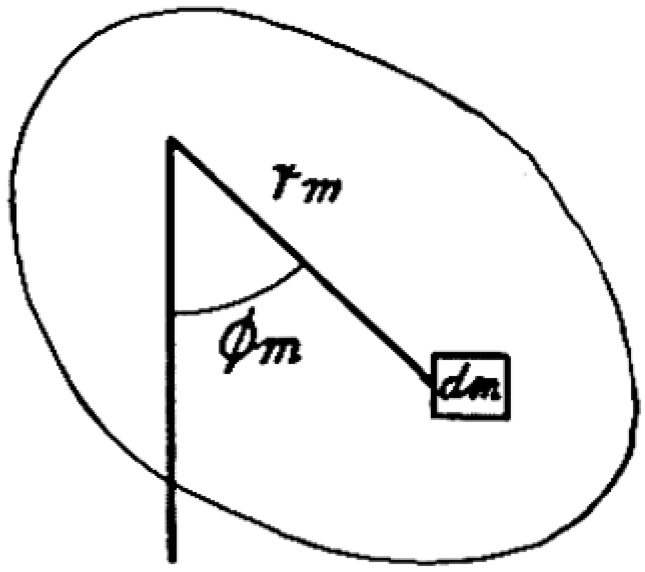


FIG. 6(*b*).

$$-\Sigma dm\, r_m{}^2 \frac{d^2\theta}{dt^2} = g\Sigma dm\, r_m \sin(\phi_m + \theta)$$

$$= g\Sigma dm\, r_m \sin\phi_m \cos\theta + g\Sigma dm\, r_m \cos\phi_m \sin\theta.$$

The first term on the right-hand side vanishes by equation (28) so that the equation of motion may be written:

$$I\frac{d^2\theta}{dt^2} + Mgh\sin\theta = 0 \quad \dots\dots\dots\dots\dots\dots \quad (29)$$

Where $I = \Sigma r_m{}^2 dm$, $M = \Sigma dm$ (the total mass of the body), and $Mh = \Sigma r_m \cos\phi_m dm$, all three of which quantities are independent of θ. Thus, for the purpose of calculating the total *moment* about the axis of the weights of the elementary masses, we may consider the whole mass of the body to be concentrated at a point G, known as *the centre of gravity*, which in equilibrium lies vertically below the axis of rotation at a distance h, while for the purpose of calculating *angular momentum about the axis* we can consider the whole mass of the body to be concentrated at a distance k from the axis, where $Mk^2 = I$, and k is known as *the radius of gyration* of the body about the axis.

Position of the centre of gravity.—We now examine the effect on I and h of displacing the axis of rotation. We first prove that *the position of G is independent of the position of this axis.* For the definition of h shows that

$$h = \frac{\Sigma r_m \cos \phi_m dm}{\Sigma dm},$$

and ϕ_m is the angle between the directions of r_m and h—that is $r_m \cos \phi_m$ is the projection of r_m on h. If h is taken along the x coordinate axis, we have

$$\bar{x} = \frac{\Sigma x_m dm}{\Sigma dm},$$

where $\bar{x}$ is the x-coordinate of G. Similar relations may be found for $\bar{y}$ and $\bar{z}$, that is, *the coordinates of G are the mean coordinates of all the elementary masses making up the body.*

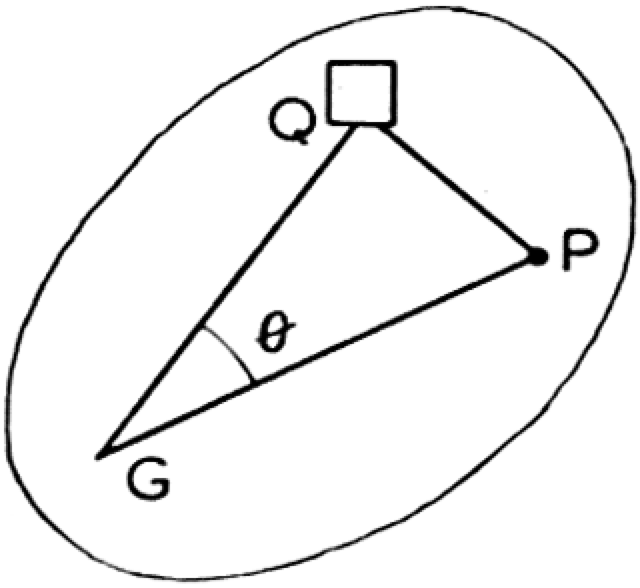


FIG. 7.

The effect on the moment of inertia of displacing the axis of rotation to a parallel position.—Let Fig. 7 represent a section of a body, and suppose that its moment of inertia through an axis through the centre of gravity G, perpendicular to the paper, is I, and we desire to know the moment of inertia I_p about a parallel axis through P.

Let Q be a typical element of mass dm.

Then, by definition,
$$\begin{aligned} I_p &= \Sigma dm\,(PQ)^2 \\ &= \Sigma dm\,(GQ^2 + PG^2 - 2PG.GQ \cos \theta) \\ &= I + PG^2 \Sigma dm - 2PG \Sigma GQ \cos \theta dm, \end{aligned}$$

using the definition of I and the fact that PG is the same for all positions of dm. Since G is the centre of gravity, the mean value over the whole body of the projection of GQ on the fixed direction PG is zero, so that the third term vanishes, and we have

$$I_p = I + M.PG^2$$

a result which is known as *the principle of parallel axes.* Thus, the moments of inertia of a body about any axis can be calculated if they are known for all axes about the centre of gravity. Incidentally, we have shown that, for any particular direction of axis, I is a minimum when the axis passes through G.

The calculation of the position of G and the value of I for a given body is a problem in integral calculus. Values for many regular shapes of body can be obtained by *Routh's rule* for calculating moments of inertia about axes through the centre of gravity. This rule states that the moment of inertia I is given by M times the sum of the squares of the two semi-axes perpendicular to the axis of rotation, divided by 3, 4, or 5 for cuboidal, elliptical, and ellipsoidal bodies respectively. (Bodies of more complicated shape can usually be broken up into pieces of these shapes, each piece making its own contribution to I, the principle of parallel axes being used as necessary.)

Thus, for a cuboid of mass M and sides a, b, and c, the moment of inertia about an axis through the centre of gravity parallel to the side of length a is $\frac{1}{3}M\,(a^2 + b^2)/4$, while for a circular cylinder of mass M and radius r the value is $\frac{1}{4}M\,(2r^2)$ about the central line, while for a sphere of mass M and radius r about an axis through the centre the value is $\frac{1}{5}M\,(2r^2)$. It is to be noticed in the application of this rule that the dimension of the solid *parallel to the axis of rotation* does not appear explicitly. It is easily seen, for example with the cuboid, that if we double the dimension c we double both the mass and the moment of inertia.

A curious "hybrid" case, not really covered by Routh's rule as it stands, is that of a cylinder rotating about an axis through the centre of gravity perpendicular to the central line. If its radius is r and its length b, the moment of inertia can be shown to be

$$\frac{M}{3}\left(\frac{b}{2}\right)^2 + \frac{M}{4}r^2.$$

We return to equation (29) for the motion of the compound pendulum. If Mk^2 is the moment of inertia of the pendulum about its centre of gravity, then its moment of inertia about P is $M\,(k^2 + l^2)$, by the principle of parallel axes. The equation of motion is of the same form as that for a simple pendulum and the period for small oscillations is

$$2\pi\sqrt{\frac{k^2 + l^2}{gl}}\left(1 + \frac{a^2}{16}\right),$$

the theory of the correction for a finite amplitude a applying here in exactly the same way as it does for the simple pendulum (p. 28). For a given body k is fixed and l varies as one alters the position of the axis of suspension relative to g. The quantity $\frac{k^2 + l^2}{gl}$ is a minimum when $l = k$, so that the minimum possible period is

$$2\pi \sqrt{\frac{2k}{g}}.$$

Any period greater than this is possible, and, solving the quadratic for l, we get two values of l corresponding to each value of T, so that there are, in general, four points on any line through G, about which the body can be pivoted to give a period equal to T (though not all of these points will necessarily lie inside the body). Suppose now that we have found two such axes at different distances (l_1, l_2) from G, which make the periods equal to T. Then we have

$$T = 2\pi \sqrt{\frac{k^2 + l_1^2}{gl_1}} = 2\pi \sqrt{\frac{k^2 + l_2^2}{gl_2}}.$$

Squaring and subtracting

$$gT^2 (l_1 - l_2) = 4\pi^2 (l_1^2 - l_2^2),$$

so that

$$T = 2\pi \sqrt{\frac{(l_1 + l_2)}{g}},$$

in other words, the distance $l_1 + l_2$ is the length of the simple pendulum of period T. This result is the basis of many modern methods of measuring g (p. 55).

Kinetic energy of a body rotating about a fixed axis.—Suppose that a body of mass M has a radius of gyration k about a fixed axis and that it is rotating about this axis with an angular velocity ω. By the principle of conservation of angular momentum we have

$$Mk^2 \frac{d\omega}{dt} = C$$

where C is the total moment about the fixed axis of the forces acting on the body. Multiplying both sides by $d\theta/dt$ or ω and integrating we get

$$\tfrac{1}{2}Mk^2\omega^2 = C\theta + \text{constant}$$

that is, *the work done by the forces in rotating the body about the axis is equal to the change in the quantity* $\frac{1}{2}Mk^2\omega^2$. It is easily verified that this quantity is in fact the total kinetic energy of the body due to its rotation, for the translational velocity of a mass dm at a distance r_m from the axis is $r_m\omega$, and its kinetic energy is therefore $\frac{1}{2}dmr_m{}^2\omega^2$. Summing this over the whole body (ω being constant for all the elementary masses if the body is rigid), the total kinetic energy is

$$\tfrac{1}{2}\omega^2 \Sigma r_m{}^2 dm = \tfrac{1}{2}Mk^2\omega^2.$$

7. Simultaneous rotation and translation. The "instantaneous axis of rotation"

Any angular velocity can be replaced by an angular velocity about a parallel axis together with a translation in a direction perpendicular to the plane containing the two axes. In Fig. 8, let a body be rotating about an axis through A, perpendicular to the paper, with an angular velocity ω. This implies that the point P is moving perpendicular to AP, with a velocity $\omega . AP$. Let this velocity be represented by the line PQ. Let PQR be a triangle whose sides PQ, QR, RP are perpendicular to AP, BA, PB respectively (Fig. 8). The two triangles are similar and, by the parallelogram law of velocities, a velocity proportional to PQ and in that direction is equivalent to the sum of two velocities represented by PR and RQ, that is, to two velocities in these directions whose values are respectively $\omega . BP$ and $\omega . AB$. The velocity $\omega . AB$ is independent of the position of P and thus corresponds to a translation of the whole body in the direction RQ, perpendicular to AB, while the velocity $\omega . BP$ for every point P represents a rotation about an axis through B, with the same angular velocity ω.

Conversely, any translation in a direction perpendicular to the axis of rotation can be compounded with the angular velocity by displacing the axis of the latter parallel to itself. The new position of this axis is sometimes known as the *instantaneous axis of rotation*. The word instantaneous is used to remind one of the fact that, in general, this axis is not fixed in space, but is continually moving parallel to itself.

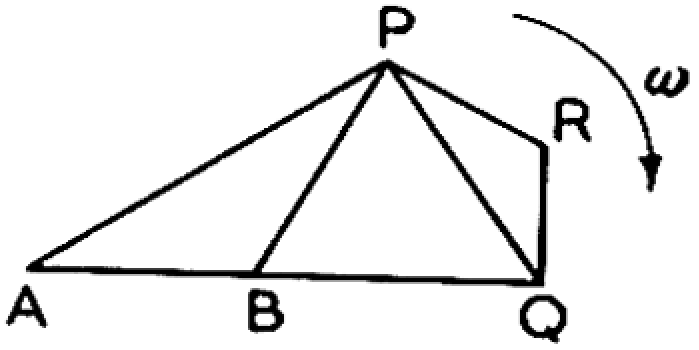


FIG. 8.

The principles of conservation of linear momentum and conservation of angular momentum were derived in Art. 3 above by arguments that are entirely *independent of one another*, so that we can always calculate the motion by equating the rate of change of linear momentum to the total force and the rate of change of angular momentum about *any* axis to the total moment of the forces about this axis. However, it sometimes happens that the linear and angular motions are related geometrically in a complicated fashion, and then it is often helpful to make use of the slightly artificial conception of the instantaneous axis of rotation. As an example where this conception is helpful, consider a circular cylinder rolling down an inclined plane [Fig. 9(a).] As the cylinder, of radius a, rolls through an angle $d\theta$, it simultaneously travels a distance $a d\theta$ down the plane. The force acting down the plane is

$$mg \sin \alpha - F$$

where F is the frictional force that is preventing sliding at P and α is the inclination of the plane to the horizontal. Thus we have

$$mg \sin \alpha - F = ma \frac{d^2\theta}{dt^2}.$$

Taking moments about the axis through the centre of gravity

$$Fa = mk^2 \frac{d^2\theta}{dt^2}.$$

Eliminating F and dividing through by m

$$g \sin \alpha = m \frac{(a^2 + k^2)}{a} \frac{d^2\theta}{dt^2} \quad \ldots\ldots\ldots\ldots\ldots \quad (30)$$

Equation (30) could, however, have been obtained immediately by equating the moment about P of the weight of the cylinder to the rate of increase of angular momentum about an axis through P, for, if the cylinder rolls through an angle θ, the line PG swings through an angle θ, so that the angular velocity about the instantaneous axis is $d\theta/dt$, while the moment of inertia about this axis is $m(a^2 + k^2)$ by the principle of parallel axes. The method can be used also, even if the cylinder is not circular, while if we handled this case by the direct method we should have to consider the linear motion both along the plane and perpendicular to it, as well as the angular rotation, thus obtaining three equations of motion from which the tangential and normal reactions of the plane would have to be eliminated, whereas the use of the instantaneous axis gives us at once

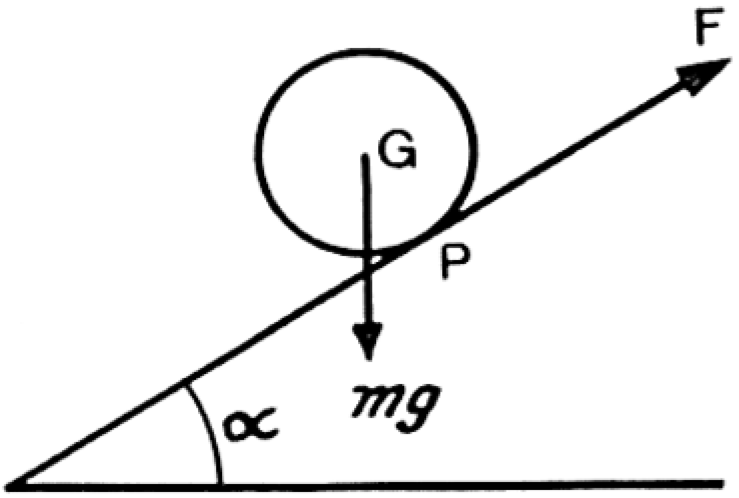


FIG. 9(a).

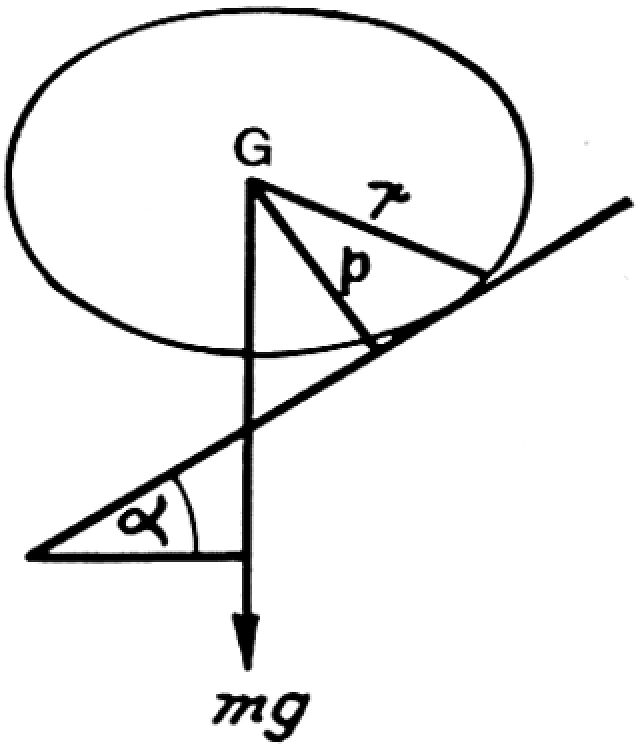


FIG. 9(b).

$$gp \sin \alpha = \frac{d}{dt}\left[(r^2 + k^2)\frac{d\theta}{dt}\right] \quad \ldots\ldots\ldots\ldots\ldots \quad (31)$$

where k is the radius of gyration about an axis through the centre of gravity, r is the polar radius of the cross-section referred to the centre of gravity as an initial point, and p is the perpendicular from the centre of gravity on to the plane [Fig. 9(*b*)], which can be calculated if r is known as a function of θ. Note that in equation (31) we have to equate the total moment to the rate of change of angular momentum, and r changes with time as well as θ if the section is non-circular.

For a uniform circular cylinder, we have by Routh's rule

$$k^2 = \frac{a^2 + a^2}{4} = \frac{a^2}{2},$$

while for a hollow cylindrical shell

$$k = a.$$

Thus by *rolling* two cylinders of equal mass down an inclined plane we can discriminate, *e.g.* between a solid cylinder and a hollow one made of denser material, even if their masses are the same.

We could have discussed the problem of the compound pendulum by considering first the motion of the centre of gravity and then the rotation about an axis through the centre of gravity, these two motions being related by the fact that the axis of suspension has to remain fixed, but again the method is more cumbersome than that of taking moments about the axis of suspension.

8. Two or more angular velocities simultaneously present. Gyroscopic motion

An angular velocity represents rotation about a certain axis, and the conception of angular velocity as a *vector* is often used, the direction of the vector being this axis, while the magnitude of the vector is numerically equal to the angular velocity about the axis, with some suitable convention as to sign, a clockwise direction of rotation looking along the positive direction of the axis usually being reckoned as positive. A similar convention is often used for couples and angular momenta, but the representation of such quantities as vectors is much less useful than the corresponding representation of velocities, forces, etc., which can be added together by the parallelogram law. It cannot be too strongly emphasised that angular velocities are *axial* vectors, which can only be added together by the parallelogram law in special cases. For example, if we try to compound two angular velocities by a process similar to that used in proving the addition law for linear velocities (Art. 1) we run up at once against the difficulty that two rotations about non-parallel axes *do not commute with one another*—for example, a rotation through θ of the y and z coordinates about the x-axis of coordinates followed by a rotation through ϕ about the y-axis is *not* equivalent to what one obtains if the order of these two rotations is reversed. Thus, a body that simultaneously possesses angular velocities about each of two non-parallel axes is essentially different from a body with a single angular velocity, and we must now investigate the nature of these differences.

One of the best known of these differences is illustrated by the gyrostat, which consists essentially of a rotating wheel mounted in gimbals so that it is capable of motion about more than one axis. Thus, in Fig. 10, suppose that the wheel has an angular velocity ω about a horizontal axis, and is then given an angular velocity Ω about a vertical axis. If the gimbals are nearly frictionless, the wheel will acquire an angular velocity about an axis PQ at right angles to the other two axes, conditions do not become steady until the vectors representing ω and Ω are parallel, and if Ω is suddenly reversed, the wheel will describe a half-turn about the axis PQ. If the direction of ω is forcibly kept horizontal, it is found that a couple $I\Omega\omega$ has to be exerted about the axis PQ, where I is the moment of inertia of the wheel about the axis of ω. This experiment is enough to show that the analogies between linear and angular velocities cannot be pushed too far, for the force on a body in a certain direction is quite independent of the velocities and accelerations in directions at right angles to it. A similar remark about couples and angular velocities would be false.

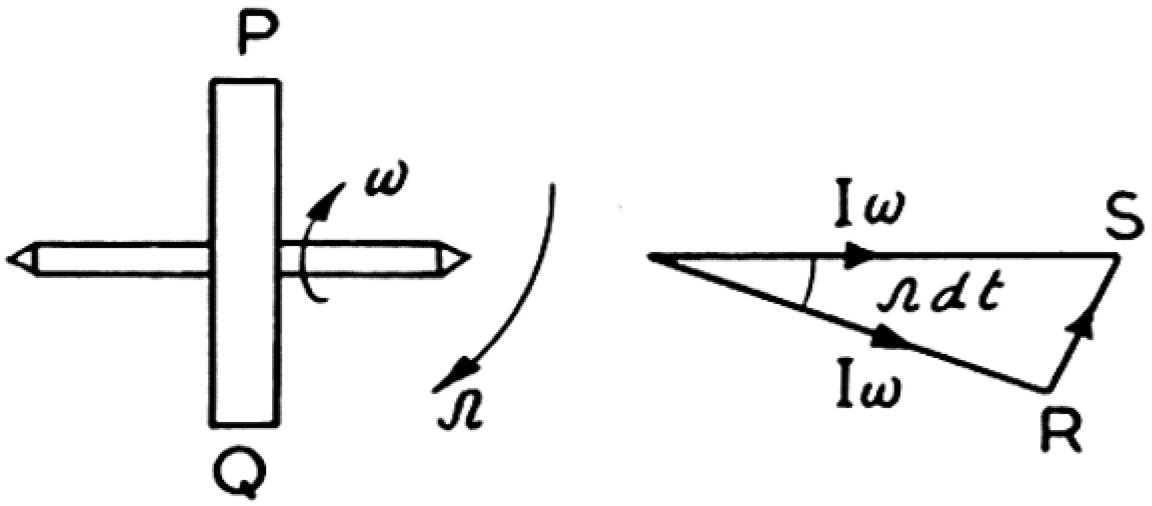


FIG. 10.

The formula for the gyrostatic couple that we have just quoted is usually proved by making the tacit assumption that angular momenta can be compounded vectorially. Thus, in Fig. 10, after a time dt, the direction of the angular momentum $I\omega$ has changed by an angle Ωdt about the vertical axis, which, according to the vector law is equivalent to a change in the angular momentum $I\Omega\omega dt$ about the axis RS, and therefore to a couple $I\Omega\omega$ about this axis. Since the vector law of addition is, in general, untrue, it is clear that we must find some better proof of this formula.

This can be done if we replace the angular velocity ω by the *equivalent distribution of linear velocities*. Suppose that the vector ω

is directed along the x-axis at a time t, and that at a slightly later time it is still in the xy plane, but now makes an angle θ with the x-axis. Now consider a point Q with coordinates X, Y, Z (Fig. 11). The linear velocity of Q, due to the angular velocity ω, is of magnitude $\omega.QR$, where QR is perpendicular to OP, and is directed along the line QS, where QS is perpendicular to both QR and OP. Let T be the point in the xy plane such that QT is perpendicular to this plane. Then the plane QRT is perpendicular to OP and therefore contains the line QS. Then RT is perpendicular to OP, and we may suppose it produced to meet QS in S. The velocity $\omega.QR$ in the direction QS may be resolved into $\omega.\frac{QR.QT}{QS}$ in the direction QT, which is parallel to the z-axis but in the negative direction, and $\omega.\frac{QR.TS}{QS}$ in the direction TS, which may be resolved into components $-\frac{\omega.QR.TS\sin\theta}{QS}$ along the x-axis, and $\frac{\omega.QR.TS}{QS}\cos\theta$ along the y-axis, since TS is perpendicular to OP, and therefore makes an angle θ with the y-axis. The angles SQR, QTR are both right angles, so the triangles SQR, STR, QTS are all similar, and we have, for the three components of velocity, due to ω in its new position

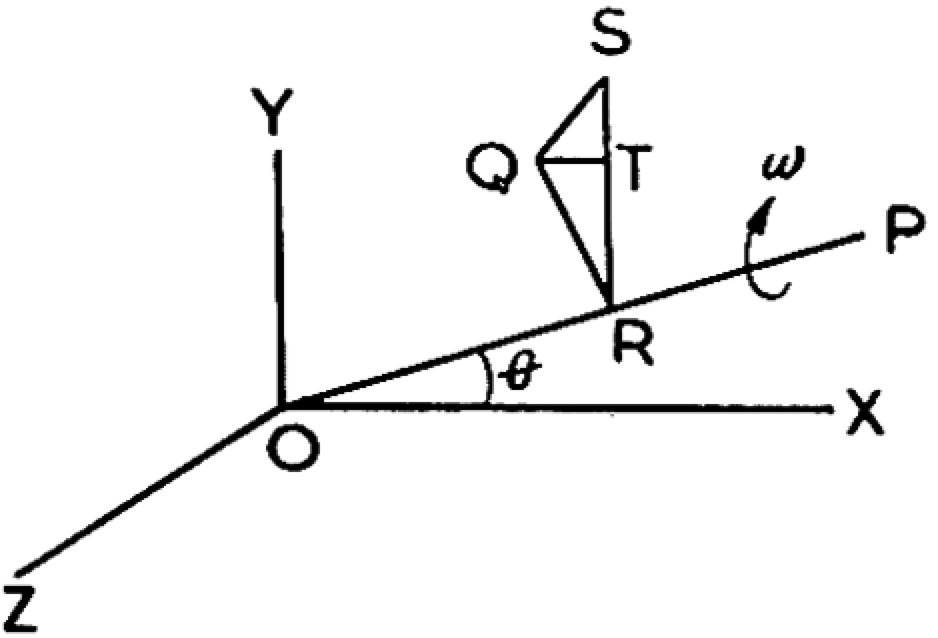


Fig. 11.

$$v_x = -\frac{\omega.QR.TS}{QS}\sin\theta = -\omega.QT\sin\theta = -\omega.Z\sin\theta$$

$$v_y = \frac{\omega.QR.TS}{QS}\cos\theta = \omega.QT\cos\theta = \omega.Z\cos\theta$$

$$v_z = -\omega.\frac{QR.QT}{QS} = -\omega.TS = \omega(X\sin\theta - Y\cos\theta).$$

Now suppose that θ is increasing uniformly with an angular velocity Ω, *i.e.* that the vector OP, representing the angular velocity ω, rotates round the z-axis with this angular velocity $d\theta/dt$ or Ω. We have

$$\frac{dv_x}{dt} = -\omega Z \cos\theta \frac{d\theta}{dt} = -\omega\Omega Z \cos\theta$$

$$\frac{dv_y}{dt} = -\omega Z \sin\theta \frac{d\theta}{dt} = -\omega\Omega Z \sin\theta$$

$$\frac{dv_z}{dt} = +\omega (X\cos\theta + Y\sin\theta)\frac{d\theta}{dt} = +\omega\Omega (X\cos\theta + Y\sin\theta).$$

For $\theta = 0$, this corresponds to a change in v_x in the time dt of $\omega\Omega Zdt$ and a change of v_z of $-\omega\Omega Xdt$, which together correspond to an angular velocity represented by a vector $\omega\Omega dt$ in the y-direction (Fig. 10) that is, just to the infinitesimal change in the vector ω due to its rotation, so that the assumption of the vector law of addition has given correct results in this particular case. It is, however, untrue that for a *finite* value of θ the effect of a rotation of the angular velocity ω, can be described by such a law, for we obtain incorrect linear velocities for points with a finite z-coordinate, as the reader can easily verify.

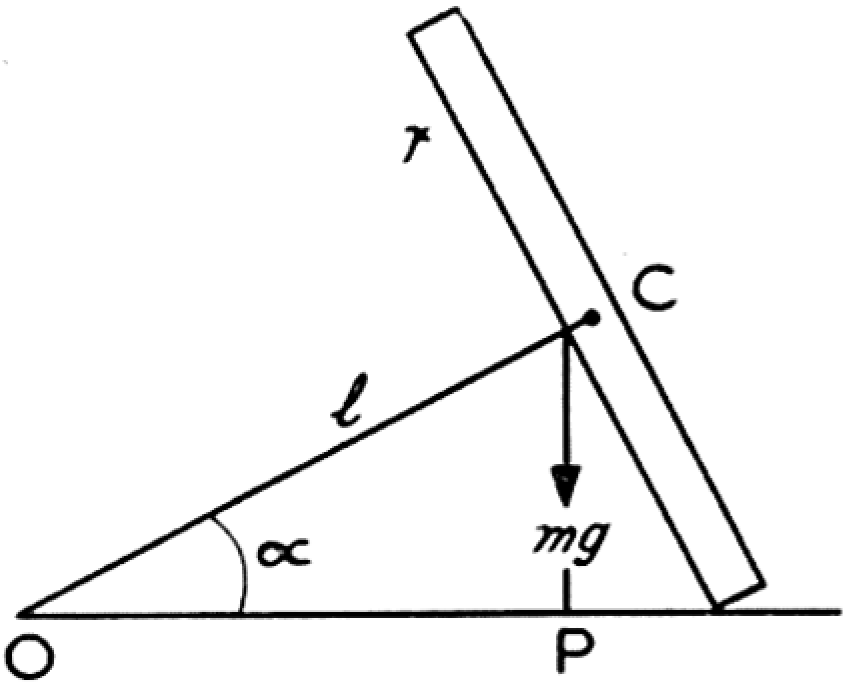


FIG. 12.

The rolling disc.—A simple problem of gyroscopic motion is that of a thin disc rolling along a plane surface. If its forward velocity is great enough, the disc is vertical, the vertical position is stable, and the path is a straight line, but as the disc slows down it leans progressively over to one side, the curvature of the path constantly increasing until the disc falls flat.

In Fig. 12, let the disc be of radius r, and let a be the "angle of lean", and let v be the velocity of its centre, and it is easily seen that O is the centre of curvature of its path on the plane. Then, the three torques acting on the disc are:

The gravitational couple $mgr \sin a$.

The "centrifugal" couple (due to the rotation about O)

$$= \frac{mv^2}{OP} . PC$$

$$= \frac{mv^2\, l \sin a}{l \cos a} = mv^2 \tan a.$$

The gyrostatic couple due to the fact that the body is rotating simultaneously with angular velocities v/r about its point of contact and v/l about $O = mk^2v^2/rl = mk^2v^2/r^2 \tan \alpha$.

For equilibrium $$\frac{k^2v^2}{r^2}\tan \alpha + v^2 \tan \alpha = gr \sin \alpha$$

which can be satisfied either by $\sin \alpha = 0$ (the upright position) or

$$v^2 = \frac{gr \cos \alpha}{1 + k^2/r^2} \text{ (the leaning position).}$$

For v^2 greater than $\dfrac{gr}{1 + k^2/r^2}$ the upright position is the only possible one for equilibrium. For v^2 less than this it is clear that the upright position is unstable, as a slight displacement from it will mean that gravity will tilt it over until α attains the value corresponding to the leaning position. For a hoop

$$k = r$$

so the critical velocity is $\sqrt{gr/2}$, while for a uniform disc

$$k^2 = \frac{r^2}{2}$$

so the critical velocity is $\sqrt{2gr/3}$.

To discuss a problem such as the behaviour of the axially-symmetrical top requires the use of coordinate axes *moving with the body* (for the various moments of inertia referred to fixed axes would vary with time), and an application of the principle of conservation of angular momentum when the various angular velocities are referred to these moving axes.

For details, the reader is referred to advanced treatises on dynamics, but the main results of physical importance will be enumerated.

(*a*) *Precession.*—As we have already seen, a body such as a gyrostat, rotating with an angular velocity ω, can remain in dynamic equilibrium with its axis of rotation horizontal and one end supported, if it has, in addition, an angular velocity Ω about a vertical axis through the supported end, such that $I\omega\Omega$ is equal to the gravitational couple. For other values of Ω equilibrium is possible with the axis inclined to the horizontal, but the problem becomes more complicated, as we have to consider as well the effect of the centrifugal force that we have just met when considering the rolling disc. In general, there are *two* possible angular velocities of precession for a given inclination of the axis and angular velocity of rotation, but one of these régimes is unstable.

(*b*) *Nutation.*—This is the name given to the oscillations of the axis of the top about its inclination of dynamic equilibrium, that occur, *e.g.* if the inclination of the axis is suddenly changed, so that oscillations about a mean value of the inclination occur. However, these oscillations are not simple harmonic, except for a small disturbance.

9. Applications of the gyroscope

We have already mentioned the rolling disc, and have seen that, provided it is rolling fast enough, a restoring moment is called into play when it leans away from the vertical. It is this effect that helps a cyclist to remain upright. It was at one time proposed to make "monorail trains", the carriages being stabilised by gyroscopes working on a similar principle. This project is technically quite feasible, and an experimental car capable of carrying a number of people was constructed, but the project was never carried any further. It is also quite feasible to diminish the rolling of a ship by gyroscopically controlled vanes. Among applications that have survived the tests of practicability and economic value may be mentioned the gyro-compass, the steering of torpedoes, and the automatic piloting of aircraft. The earth is a gyroscope, and gyroscopic effects manifest themselves in *the precession of the equinoxes.*

The gyro-compass may be regarded as typical of the applications which have the object of fixing a *direction* within a moving body. As we have seen, if a gyroscope is given an angular velocity about an axis perpendicular to its axis of spin, a couple, and consequent motion, about the *third* axis result. This motion can, as in the torpedo, be made to open valves controlling the vertical rudders so as to bring the torpedo back to its course. In the gyro-compass it is desired that the gyroscope axis itself should point parallel to the earth's axis, and that a *restoring* torque should come into play in the event of any actual deviation. This does not happen in the ordinary way; after a deviation the axis *precesses* round its proper direction. By making the gimbals of the gyroscope *bottom-heavy* it is possible to make the instrument self-righting. If the gyroscope is at the equator and spinning parallel to the earth's axis, the imposition on it of an east to west angular velocity will result in a torque tending to impart a north to south angular velocity. If the gimbals are made bottom-heavy, the consequent displacement will bring into action a couple tending to correct the original east to west deviation. Thus, this ingenious idea deals with the difficulty that, with a simple gyroscope, the torque called into play by an angular velocity is not directly a *restoring* torque, but acts about an axis *at right angles* to the axis of that angular velocity. However, the gyro-compass is subject to a number of errors, as it is not able to distinguish between the effects of accelerations due to the angular velocity of the earth and those of the acceleration, pitching, and rolling of the ship. Thus, a practical gyro-compass needs several corrections, some of which can be applied automatically by smaller gyros incorporated in the compass, while others have to be computed from tables according to the course and speed of the ship.

The *precession of the equinoxes* is due to the fact that the earth is not an exact sphere but bulges slightly at the equator, in addition to which the sun and moon are not usually in the equatorial plane of the earth, but are approximately in the plane of the ecliptic, which is inclined to the equatorial plane at an angle of about 23½°. The gravitational attraction of the sun and moon on the equatorial bulge produces a torque leading to a precession of the earth's axis, which describes a cone relative to the fixed stars, quite similar to the cone described by the axis of a precessing top as a result of the couple due to its weight. As the couple due to the sun and moon is comparatively small, it takes several thousand years for the earth's axis to describe the complete cone, but the slow drift of the polar axis relative to the fixed stars is quite easily measurable astronomically.

10. Bifilar suspension

We consider a beam suspended as in Fig. 13. This method of suspension is used in scientific instruments:

(*a*) When it is desired to magnify a small displacement. As an example of this we may mention the Kelvin doubly-suspended mirror (p. 52).

(*b*) As a simple means of determining moments of inertia without using a torsion wire.

(*c*) When one desires to obtain a controlling torque that can be *continuously varied*. This cannot be done at all conveniently with an ordinary torsion suspension.

In Fig. 13, let l be the length of each thread, $2b$ the distance between the tops of the suspensions, $2a$ the distance between the bottoms of the suspensions, M the mass of the beam, and k its radius of gyration about a vertical axis through the centre. Suppose the beam twisted through an angle θ in a horizontal plane, so that the bottom end of one string is at A_1 instead of A. Let A_2 be the point vertically above A_1 and in the same horizontal plane as B, and let B_1 be the point on OA vertically below B. Let ϕ be the angle BA_1A_2 at which each string is inclined to the vertical. Then, resolving vertically,

$$2T\cos\phi = Mg \quad \dots\dots\dots\dots\dots\dots \quad (32)$$

where T is the tension in each string.

The horizontal component of the tension of each string is $2T\sin\phi$ and lies along the line A_1B_1. The component of this along the beam has no effect, because it is balanced by a similar force due to the other string. The component perpendicular to the beam is $T \sin \phi \sin \alpha$, where α is the angle OA_1B_1, so that the total restoring moment due to the two strings is $2Ta \sin \phi \sin \alpha$, and the problem is solved if we can determine ϕ and α in terms of θ, which can be done from the geometry of the triangles OA_1B_1, A_1BB_1. From triangle OA_1B_1 we have

$$A_1B_1 = (a^2 + b^2 - 2ab \cos \theta)^{\frac{1}{2}} \quad \ldots\ldots\ldots\ldots \quad (33)$$

and

$$\frac{A_1B_1}{\sin \theta} = \frac{b}{\sin \alpha} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (34)$$

From triangle A_1BB_1,

$$\cos \phi = \frac{l^2 + BB_1{}^2 - A_1B_1{}^2}{2l \,.\, BB_1} \quad \ldots\ldots\ldots\ldots \quad (35)$$

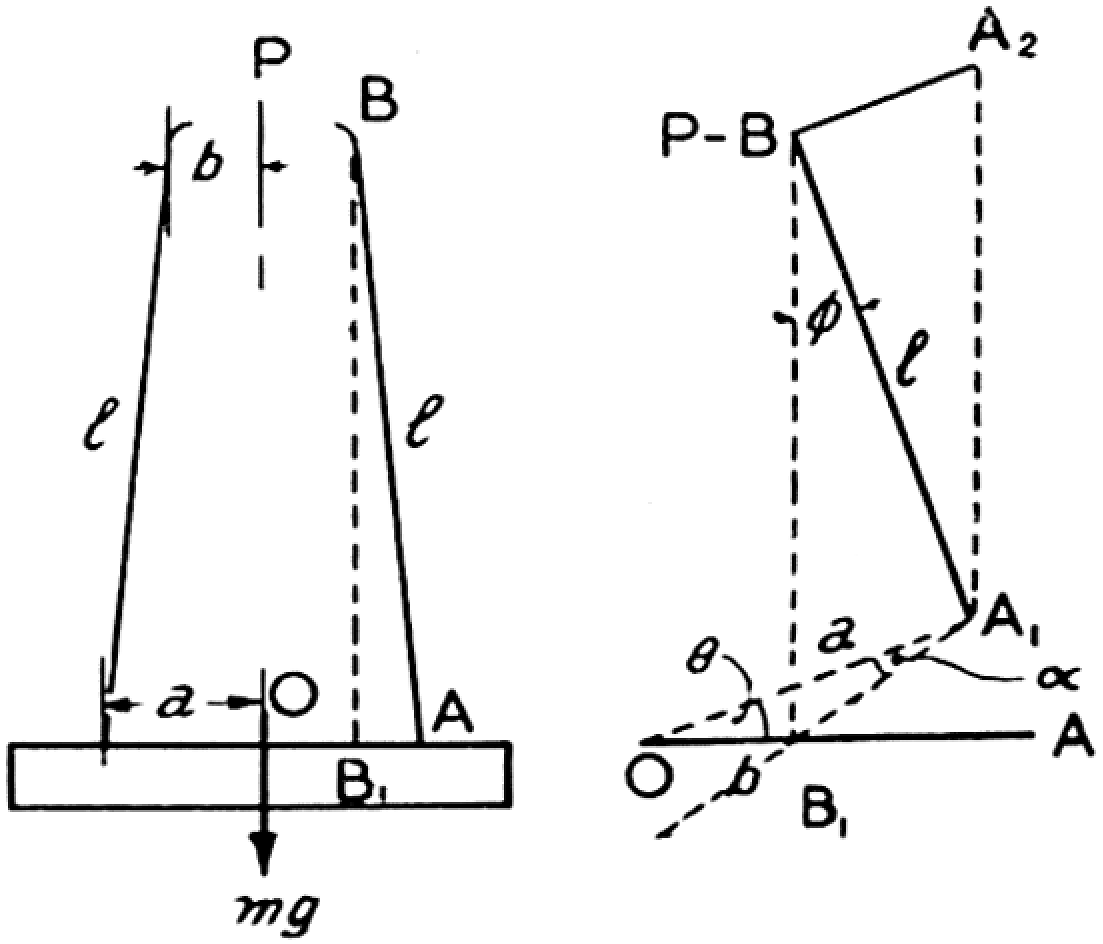


FIG. 13.

But BB_1 is the depth of the beam below the supports in the position of equilibrium, so that $BB_1{}^2 = l^2 + (a - b)^2$.

These formulae are all exact, but the equation of motion is extremely cumbersome. It takes a much simpler form if we assume that θ is small, so that we may replace $\sin \theta$ by θ, and $\cos \theta$ by 1, as in the theory of the simple pendulum. We then have

from (33) $$A_1B_1 \simeq |a - b| = x, \text{ say} \quad \ldots\ldots\ldots \quad (36)$$

from (34) $$\frac{\alpha}{\theta} \simeq \frac{\sin \alpha}{\sin \theta} = \frac{b}{A_1B_1} \simeq \frac{b}{|a - b|} \quad \ldots\ldots \quad (37)$$

from (35) $$\sin \phi \simeq \frac{|a - b|}{l} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (38)$$

So that the equation of motion is

$$
\begin{aligned}
-Mk^2\frac{d^2\theta}{dt^2} &= 2Ta\sin\phi\sin\alpha \\
&= Mga\tan\phi\sin\alpha \qquad \text{from (32)} \\
&\simeq Mga\,\frac{x}{(l^2-x^2)^{\frac{1}{2}}}\cdot\frac{b}{x}\cdot\theta \qquad \text{from (37) and (38)}
\end{aligned}
$$

so that the motion is simple harmonic with period

$$2\pi\sqrt{\frac{k^2\,(l^2-x^2)^{\frac{1}{2}}}{g\,ab}}$$

and the length of the simple equivalent pendulum is

$$\left(\frac{k^2\,(l^2-x^2)^{\frac{1}{2}}}{ab}\right)$$

and it can be seen how k can be determined from a measurement of the period of the oscillation, in terms of g and easily measurable lengths. If the strings are vertical $x = 0$ and $a = b$, so that the length of the simple equivalent pendulum is k^2l/a^2. If either a or b is made small, the period can be made very long and the controlling couple very small. In practice, it is easier to make b small than it is a.

11. Conclusion

Although the fundamental laws of mechanics are quite simple, they are not directly applicable to actual problems, and, by considering a number of typical problems that occur in physics, we have attempted to bring out the intermediate steps. The reader is referred to textbooks of mechanics for further details of the various methods. Many of the problems we have considered could have been handled by one general method, the Lagrangian *method of generalised coordinates*. However, this method, although of very great beauty and mathematical simplicity when the forces are *conservative*, becomes very difficult to apply when forces such as friction are involved. Although the method is of fundamental significance in quantum mechanics and electromagnetic theory, its use has been avoided in this book.

BIBLIOGRAPHY FOR CHAPTER II

S. L. Loney, *Statics*. C.U.P. (1930).

A. S. Ramsey, *Dynamics*. 2nd Edn. C.U.P. (1933-7).

E. J. Routh, *Advanced Rigid Dynamics*. 6th Edn. Macmillan (1905).

Examples on Chapter II

1. Summarise as concisely as possible the non-astronomical evidence that leads us to believe in Newton's laws of motion.

2. In Fig. 2 derive the tension in the string GH without using the principle of virtual work. (Hint.—Consider the equilibrium of the systems $ABCDEF$, $BCDE$, CD in succession, *e.g.* the external forces acting on CD are its weight, the reactions at C and D, and the tension T.)

3. In Fig. 3 use the principle of virtual work to determine the condition of equilibrium if the distance from A to B or D to E is 1·01 times that from B to C or E to F, and AC, DE are supposed to be uniform rods each of mass m_3.

4. Determine the equation of the curve which touches all possible trajectories of a particle projected from the origin under gravity, with an initial velocity v_0. Show also that it encloses all such trajectories.

5. Determine the moment of inertia of a right circular cone of mass M, height h, and semi-vertical angle θ about its axis, if it be supposed made of material of uniform density ρ.

6. Determine the critical velocity at which a coin of radius 1 cm. can roll along a level surface while remaining upright.

7. A pile of mass M is subject to a resisting force p which is taken as constant. The pile-head of mass m strikes it with velocity v and does not rebound. Determine the subsequent motion until the pile is brought to rest:

(*a*) On the assumption that the force between pile and head is a large constant force acting for a very short time, afterwards dropping to zero when the velocities become equal.

(*b*) By making use of the principle of conservation of momentum.

Explain clearly why these two approaches give the same final result, in the limiting case when the time interval is very short. Determine the fraction of the kinetic energy of the pile-head that is used in doing work against the force p. What becomes of the remaining energy?

8. A hollow circular cylinder, made of uniform material with an outer radius of twice the inner radius, is rolled down an inclined plane. What is the greatest angle at which the plane may be inclined to the horizon without slipping taking place? Take coefficient of friction as $\frac{1}{2}$.

9. Compare the velocities at aphelion and perihelion of a comet, the major and minor axes of the orbit being a and b. (Hint.—Use rule of equal areas.)

10. Neglecting atmospheric resistance, determine the minimum velocity at which a body must be projected in order that

(*a*) It may become a satellite of the earth.

(*b*) It may leave the neighbourhood of the earth altogether.

In each case, determine the angle of projection. (Take the acceleration of gravity at the surface of the earth as 980 cm./sec.^{2} and assume an inverse-square law of force.)

CHAPTER III

GRAVITATION

1. Newton's Law of Gravitation

The law of gravitation formulated by **Newton** states that the attraction exerted by any small portion of matter on another depends only on their masses and their distance apart. If one is dealing with bodies of finite size, the total force between them is the resultant of the individual actions of the small particles of which we may suppose the bodies to be composed. The law assumed by Newton is that the attraction is proportional to the inverse square of the distance between any two particles, and that it is also proportional to the product of their masses. This may be expressed as follows:

$$F = G\frac{m_1 m_2}{d^2}$$

where G is a universal constant. This law is now known to be not quite correct, but it is an extremely close approximation and there are only a few phenomena that need for their description the more exact law formulated by **Einstein**. We shall first develop some consequences of Newton's law and then briefly describe these exceptional phenomena.

Newton's law is of inverse-square type, as are the laws of electrostatics and magnetostatics, so that much of the mathematical theory of gravitation is identical with that of electrostatics. There are, however, some very important physical differences.

(*a*) *Absence of any screening effects.*—As far as is known experimentally, there is nothing corresponding to the *dielectric constant* of a medium, the gravitational force between two bodies being quite unaffected by, *e.g.* evacuating the region between them or altering the medium. As we shall see later, the period of oscillation of a pendulum can be determined with very high accuracy, and a change in that period of the order of 1 in 10^6 could be detected with certainty, so that "gravitational permeabilities" differing from unity by as little as this could have been detected.

(*b*) *Absence of directional effects.*—Since so many of the physical properties of crystals, in particular their dielectric constants, vary with their orientation, it is natural to ask whether the mutual orientations of, *e.g.*, two quartz crystals has any effect on their gravitational attraction. Such an effect was looked for by **Poynting** and **Gray** by revolving a quartz sphere close to another suspended sphere. If directional effects were present, a periodic couple on the suspended sphere would result, and a large oscillation of the latter would occur if the torsional period of the suspended sphere were nearly equal to the period of this couple. No such effect could be found.

(*c*) *Absence of temperature effects.*—No positive evidence of any change in gravitational attraction as the attracting bodies are heated has been obtained in terrestrial experiments, and astronomical evidence, which seems to show that the constant in Newton's law is the same for bodies at all kinds of temperatures, makes such an effect unlikely.

The theory of relativity makes it likely that "gravitational waves" can be propagated from a moving body just as electromagnetic waves are propagated from a moving charge, and that the velocity of propagation is the velocity of light *in vacuo,* but this prediction has never been verified or disproved experimentally on account of the high velocity of propagation and the numerical smallness of the expected amplitudes.

2. Gravitational field and gravitational potential

The existence of a law of force proportional to mass enables one to define a gravitational field quite analogous to the electric field. The *gravitational field* at a given point is the force that would act on unit mass placed at that point. This is a vector quantity, and it is often more convenient to calculate the *gravitational potential*, the gravitational potential of a body, or set of bodies, at a point being the work done by gravitational attraction when a unit mass is brought up to this point from a great distance. From the principle of conservation of energy it follows that the work thus done must be independent of the path followed by the test-mass, because it would otherwise be possible to gain energy by continually bringing up the test-mass along one path and withdrawing it along another.

We now prove that the gravitational potential at a point due to any number of small masses is equal to the sum of the potentials that each would produce if present separately. In Fig. 14, let m_1, m_2 be two small masses with coordinates (x_1, y_1, z_1), (x_2, y_2, z_2). The force on unit mass at point (X, Y, Z), due to m_1, is by Newton's law,

$$\frac{Gm_1}{(X - x_1)^2 + (Y - y_1)^2 + (Z - z_1)^2}$$

directed along the line Pm_1, and the x-component of this force is therefore

$$\frac{Gm_1}{(X-x_1)^2+(Y-y_1)^2+(Z-z_1)^2}\cdot\frac{x_1-X}{[(X-x_1)^2+(Y-y_1)^2+(Z-z_1)^2]^{\frac{1}{2}}}$$

since $x_1 - X$ is the resolved part of the line Pm_1 along the x-axis. This force is equal to

$$\frac{\partial}{\partial X}\,\frac{Gm_1}{[(X-x_1)^2+(Y-y_1)^2+(Z-z_1)^2]^{\frac{1}{2}}}$$

and similar expressions apply for the y and z components of the force.

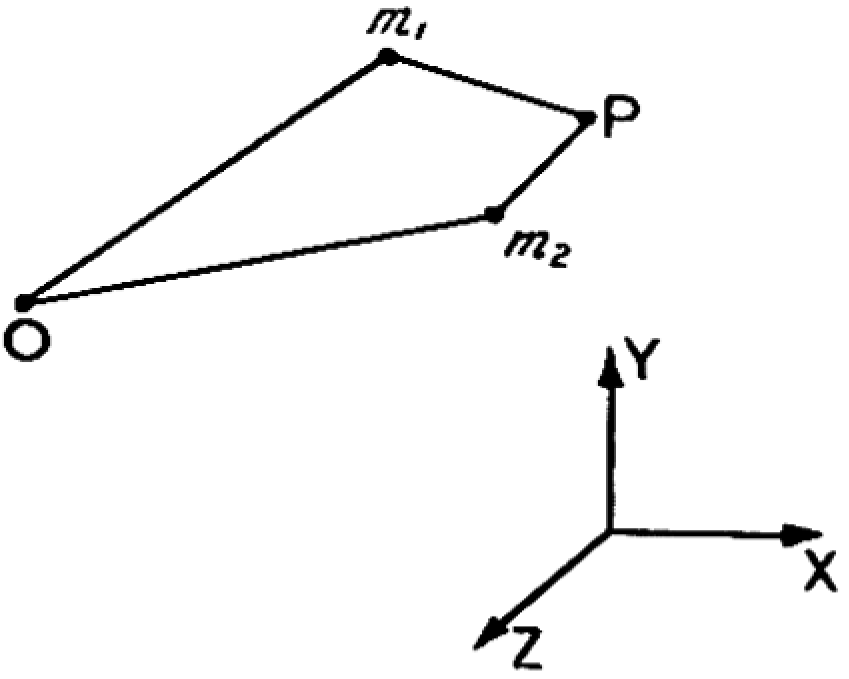


FIG. 14.

Thus we can calculate the forces due to m_1 and m_2 together if we know the quantity

$$-\frac{Gm_1}{[(X-x_1)^2+(Y-y_1)^2+(Z-z_1)^2]^{\frac{1}{2}}}$$
$$-\frac{Gm_2}{[(X-x_2)^2+(Y-y_2)^2+(Z-z_2)^2]^{\frac{1}{2}}}$$

and this result can be extended to any number of small masses, or to a large body regarded as an aggregate of very small masses. Further, the x-component of the force on a body is the rate of change of energy with distance as that body is displaced in the x-direction, so that we have identified the quantity

$$-G\sum_i \frac{m_i}{r_i}$$

where r_i is the distance from the mass m_i to the point in question, as the gravitational potential, and the gravitational field can be calculated if the potential is known. The negative sign is used to comply with the sign convention relating to potential energy—the doing of work by the gravitational force implies a decrease in the associated potential energy. We now consider a few special cases.

(*a*) *The potential of a thin spherical shell.*—In Fig. 15 P is a point at a distance r from the centre of a thin spherical shell of mass M, of radius R and wall thickness t. Consider thin circular rings with centres on OP, of which the typical one subtends an angle θ at O and has width $R\delta\theta$. Each element of this ring is at the same distance, $PQ = x$, from P. The contribution of the ring to the potential at P is

$$-\frac{G.2\pi R\sin\theta.R\delta\theta.t.m}{x}$$

where m is the density of the shell and the total potential is obtained

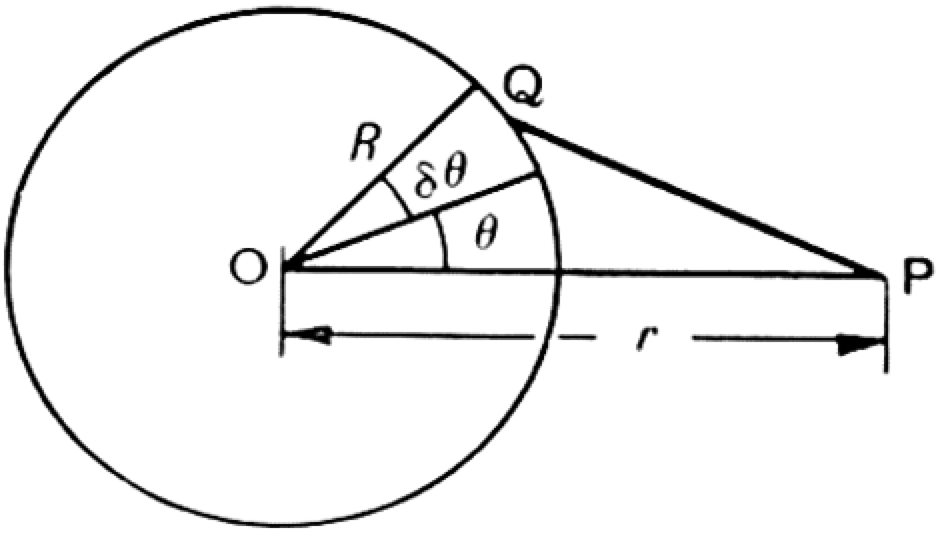


FIG. 15.

by integrating over all values of θ between 0 and π, using the relation

$$x^2 = R^2 + r^2 - 2Rr\cos\theta$$

Thus $$V = -2\pi R^2Gmt\int_0^\pi \frac{\sin\theta d\theta}{(R^2 + r^2 - 2Rr\cos\theta)^{\frac{1}{2}}}.$$

Changing the variable to $y = \cos\theta$, we have

$$dy = -\sin\theta d\theta$$

and the limits of integration become $+1$ and -1, so that

$$V = -2\pi R^2Gmt\int_{-1}^{+1}\frac{dy}{(R^2 + r^2 - 2Rry)^{\frac{1}{2}}}$$

$$= +2\pi R^2Gmt\,{}_{-1}^{+1}\left[\frac{(R^2 + r^2 - 2Rry)^{\frac{1}{2}}}{Rr}\right].$$

A difficulty arises here because it is not obvious which signs to use when taking the square roots, but, since the quantity $(R^2 + r^2 - 2Rry)^{\frac{1}{2}}$ represents the distance from P to a typical point on the sphere, it follows that the positive square roots are to be taken, in all cases.

Thus for $r > R$
$$V = \frac{2\pi R^2 Gmt}{Rr}[(r - R) - (R + r)]$$
$$= -\frac{4\pi R^2 Gmt}{r}$$

for $r < R$
$$V = \frac{2\pi R^2 Gmt}{Rr}[(R - r) - (R + r)]$$
$$= -\frac{4\pi R^2 Gmt}{R}$$

the second formula also holding for $r = R$. Since the mass of the shell M is equal to $4\pi R^2 mt$, the potential outside the shell is $-MG/r$, while inside the shell it remains constant. Thus, outside the shell, the gravitational potential (and therefore the gravitational field also) is the same as one would get from a mass M at the centre of the sphere, while inside the shell the potential is constant, so that the field vanishes. The latter result is equivalent to the celebrated result in electrostatics that a hollow charged sphere exerts no force on a charge inside it, and such results hold only for an inverse-square law of force.

(*b*) *The potential of a solid sphere.*—The potential and field of a solid sphere can at once be inferred from those of a spherical shell, even if the density varies with the distance from the centre. Inside the sphere, all the matter at a greater distance from the centre than the point in question contributes nothing to the field, so that the attraction is that due to the matter inside a concentric sphere passing through the point, while, outside the sphere, the attraction of all the shells is the same as if their masses were concentrated at the centre. This result is of considerable historic importance, as it enabled **Newton** to apply his law of gravitation to the moon's motion. The radius of the earth is not negligible compared with that of the moon's orbit, so that, without this result, there would have been no means of deciding whether the distribution of matter inside the earth, and its finite radius, necessitated any correcting terms in the equation of motion of the moon.

For more complicated cases, *e.g.* for an ellipsoidal body, or for a density of matter varying with polar angle as well as with radius, the reader is referred to treatises on potential theory.

3. Inertial mass and gravitational mass

Newton's law of gravitation implies the very important result that the gravitational field *produced by a body* is proportional to its mass. We defined *mass* (p. 7) by means of the *acceleration produced on it* by a known force, and the mass of a body defined in this way is *inertial mass.* It is, however, equally possible to define the mass of a body as proportional to the force produced *by its gravitational attraction on a standard test body* at unit distance, and the mass defined in this way is called the *gravitational mass.* **Galileo** found that the *acceleration of falling bodies was independent of their mass*, and a similar result has been found to hold with great accuracy for pendulums used in measuring the acceleration of gravity (p. 54), showing that *the gravitational force between a mass and the earth is proportional to the inertia of that mass.* There seems to be no *a priori* reason for expecting such a result [if, for example, one considers a body *rolling* down an inclined plane, the acceleration depends not only on the mass but also on the distribution of that mass in space (p. 35), so that it must be regarded as an experimental law]. It is sometimes claimed that the equality of inertial and gravitational mass is *predicted* by the general theory of relativity, but opinions differ on this point.

The equality of inertial and gravitational mass (we have shown that they are proportional, so that they can obviously be made equal by a proper choice of units) has very important consequences in astronomy. For example, the orbit of a satellite round a planet is *independent of the mass of the satellite* because of the proportionality of gravitational force and inertial mass, so that observations of the orbit of the satellite enable one to "weigh" the planet—similarly, the masses of *each* component of a double star can be calculated from the orbit of the other component round the common centre of gravity, since the constant G can be found from *terrestrial* experiments.

4. Measurements of the gravitational constant G

Since the acceleration of gravity at the earth's surface can be measured very accurately (p. 54), a measurement of G enables us at once to infer a value of M, the mass of the earth, so that such experiments are sometimes described as "weighing the earth". The experiments are all based on either the direct laboratory measurement of the force between two masses, or on the *comparison* of the force exerted by a known mass on another mass with the attraction of the earth on the second mass, and one may distinguish them roughly as measurements of G and measurements of M respectively.

(*a*) *Measurements of M.*—The first attempt was due to **Bouguer,** who attempted to measure the effect of the attraction of a mountain (Chimborazo) on the direction of a plumb-line. By taking astronomical observations at a station near the mountain and at a second station some miles away, the shift in the direction of a plumb-line due to the attraction of the mountain on the pendulum bob could be measured.

Various other attempts to use a mountain in this way have been made and the method enables the order of magnitude of M to be determined. An experiment on a similar principle is due to **Airy**, namely, a measurement by a pendulum method (p. 54) of the variation in the acceleration of gravity between the surface of the earth and the bottom of a mine. The earth is so nearly spherical that we can apply the theory outlined above (p. 49) which shows that the spherical shell of the earth which is further away from the earth's centre than the observer contributes nothing to the force of gravity. We thus have, if g_1 is the acceleration of gravity at the earth's surface and g_2 the acceleration at the bottom of a mine of depth h, M the mass of the earth, R its radius, and ρ the mean density of the surface crust, the results

$$g_1 = \frac{GM}{R^2} \qquad g_2 = \frac{G\,(M - 4\pi R^2 \rho h)}{(R-h)^2}$$

so that if g_1/g_2 is measured, M is known if ρ is known.

One difficulty in all experiments of this type is that of getting a satisfactory estimate of the mass of the body with which M is being compared, *i.e.* the mass of the mountain, or the mass of a surface layer of the earth of thickness h. In addition, the mountain experiments suffer from a more fundamental difficulty to be discussed later (p. 67). One type of experiment that has been proposed is to measure the deflection of a plumb-line near a tidal estuary as the tide rises and falls. Here the mass of the additional water can be determined accurately, but one cannot be sure that the weight of this water is not pressing down the bottom of the estuary and thus causing an additional disturbance to the plumb-line. A better method was suggested by **von Jolly**, that of placing a large mass under one of the arms of a balance and thus measuring directly the pull of this mass on the mass in the pan. Von Jolly did actually carry out such an experiment, and obtained fairly good results, but the experiment is usually known as **Poynting's** experiment as he carried it out under carefully controlled conditions and obtained a really reliable result.

Poynting's apparatus is shown diagrammatically in Fig. 16(*a*). A sensitive bullion-balance was placed in an underground room, and all manipulations were made by rods controlled from outside, to avoid trouble due to air currents. The masses A and B were lead spheres of about 50 lb. each. The deflection of the pointer due to the movement of a rider along the beam was first determined, and compared with that caused by introducing a large lead sphere M weighing about 350 lb., under one of the masses. It was found that the presence of M_1 weighed down the floor and caused an apparent shift in the vertical, and the mass M_2, about half as big as M_1, was placed with it on a turntable at about twice the distance of M_1 from the pivot T_1, as shown in the figure, so that the turntable could be rotated to bring M_1 successively under A and B without tilting the floor. The effect of M_1 on the more distant of the two masses had to be corrected for, likewise, the effect of M_2 on both masses. The effect of M_1 and M_2 on the beam of the balance was difficult to calculate, but it was eliminated by repeating the experiment with the two masses A and B suspended about a foot higher. The movements of the pointer were small and were magnified by the Kelvin double-suspension mirror arrangement shown in Fig. 16(*b*). The motion of the mirror was observed by the usual beam of light and scale and an additional magnifying effect could be obtained when required if the distance CD was made small. Since what is actually wanted is the *ratio* of the deflection produced by a rider of known mass to that produced by the attractive force, all calibration of the double-suspension arrangement can be avoided provided that both parts of the experiment are carried out with the same setting of CD. A dash-pot filled with oil was used to help damp out the oscillations of the mirror.

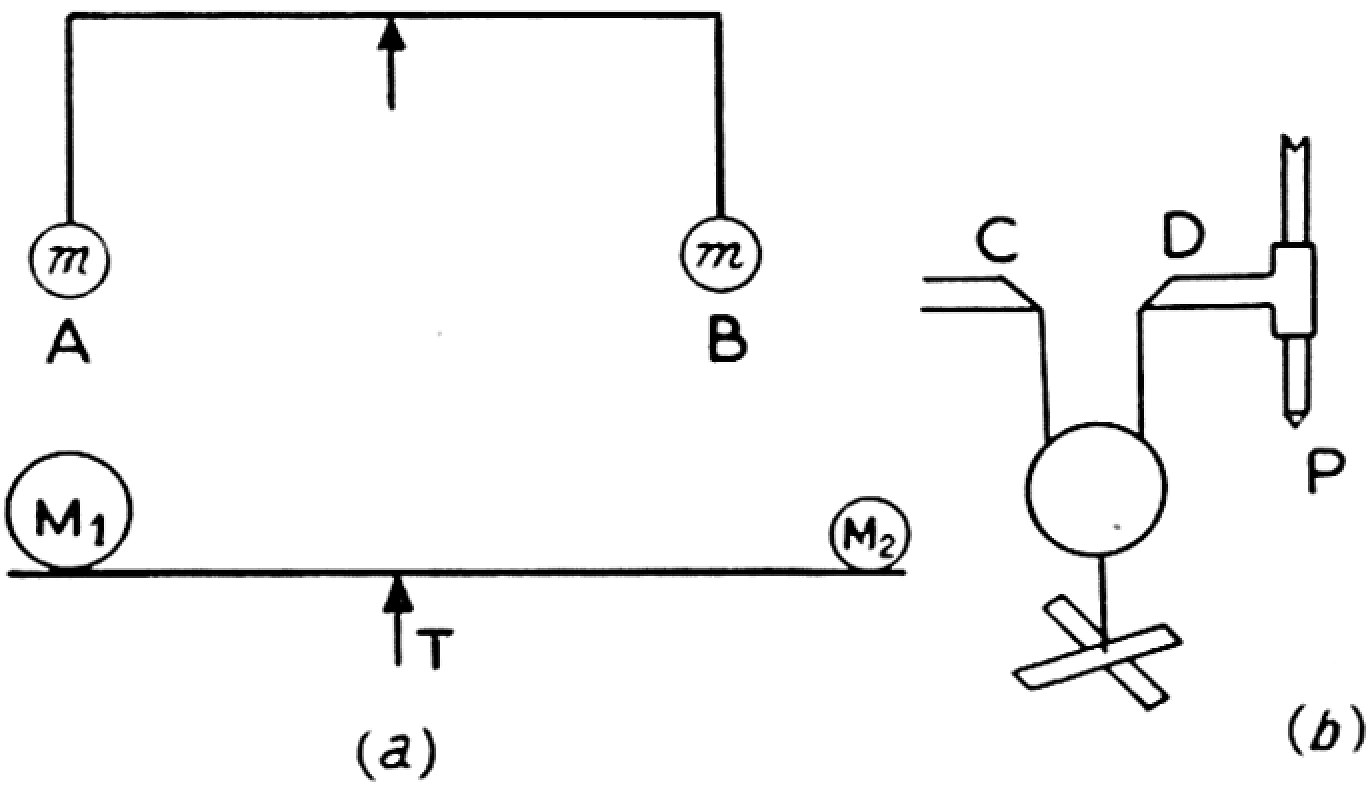


FIG. 16.

(*b*) *Direct measurements of G. The Cavendish-Boys experiment.*—The first determination of M or G using masses of laboratory size was due to **Cavendish**. His apparatus consisted simply of a torsion balance, a long rod carrying massive spheres at each end and suspended in the middle by a torsion wire, and the experiment involved measuring the displacement of the position of equilibrium when two large spheres were brought near the ends of the beam on opposite sides. As the apparatus used was rather large, the beam being 6 ft long, it was housed in a special room, and manipulated from outside, but air-currents caused a great deal of trouble, and the positions of equilibrium had to be determined by observing the swings of the beam by means of a telescope, each end of the beam carrying a vernier moving over a fixed scale. Despite the experimental difficulties, Cavendish obtained a value differing by less than one-half per cent. from the present accepted value.

An improvement in the apparatus was made by **Boys,** who was able to diminish the size of the apparatus considerably without loss of sensitivity by using as suspensions the extremely fine quartz fibres, the process of making which he discovered. These combine great tensile strength with small resistance to torsion, so that the apparatus could be made small and yet reasonably sized attracted masses could be used. The apparatus is illustrated diagrammatically in Fig. 17. The apparatus was totally enclosed to avoid air-currents. The beam AB also acted as a mirror, enabling the deflection to be measured optically. The two small masses m, were gold spheres, about ¼ in. in diameter, and were mounted at different heights in order to diminish the opposing gravitational couple due to the attraction of the more distant lead sphere on each gold sphere, which would become a big fraction of the couple to be observed if a short torsion arm were used with the masses at the same level. The attracting masses were lead spheres about 4½ in. in diameter.

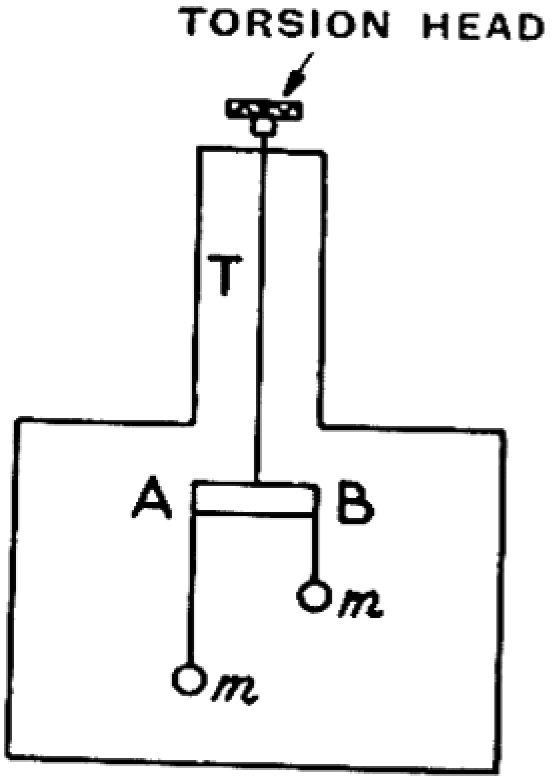


Fig. 17.

A review of all the measurements of G and M up to 1901 was made by **Burgess,** who considered that the most reliable value of G was given by

$$G = (6{\cdot}6607 \pm 0{\cdot}0016) \times 10^{-8} \text{ C.G.S. units.}$$

This differs by less than one part in 200 from Cavendish's value, and by less than one part in 2,000 from Boys' value. The corresponding value of the mean density of the earth is

$$D = 5{\cdot}5247 \pm 0{\cdot}0013 \text{ g. per c.c.}$$

A number of later experiments giving values practically identical with this have been made. **Zahradnicek** used a resonance method, an oscillating heavy torsion balance with large lead spheres mounted on the ends of a bridge-shaped beam, inducing oscillations in an independently-suspended torsion balance having a light beam with spherical masses at the ends. **Heyl** carried out an experiment on what may be thought of almost as a modernised version of Cavendish's original apparatus. He investigated the change in *period* of a torsion balance, consisting of two masses at the ends of a beam suspended from a tungsten filament, that occurred when two large attracting masses were swung into position near each end of the beam and in line with it. A change in the *period* can be measured much more easily and accurately than the change in the *zero* of the torsion balance, due to the presence of attracting masses, which is what Cavendish measured.

5. Measurements of g, the acceleration due to gravity and its variations

Accurate values of the acceleration of gravity are needed by physicists for two quite distinct reasons. First, it is needed whenever a physical quantity involving a *force*, such as the barometric pressure or Young's modulus of a wire, is measured. If we do not know g we can only measure such quantities in grams-weight per square centimetre, whereas, to get a valid comparison with work in other laboratories such quantities must be expressed in dynes per square centimetre. Secondly, measurements of the variation of g from place to place gives information on the shape of the earth and its structure. The absolute measurements of g depend on pendulum experiments, which are difficult on account of the many corrections involved, and various simpler instruments have been devised for measuring the *difference* of g between two stations. It is also possible to measure the *gradient* of g directly.

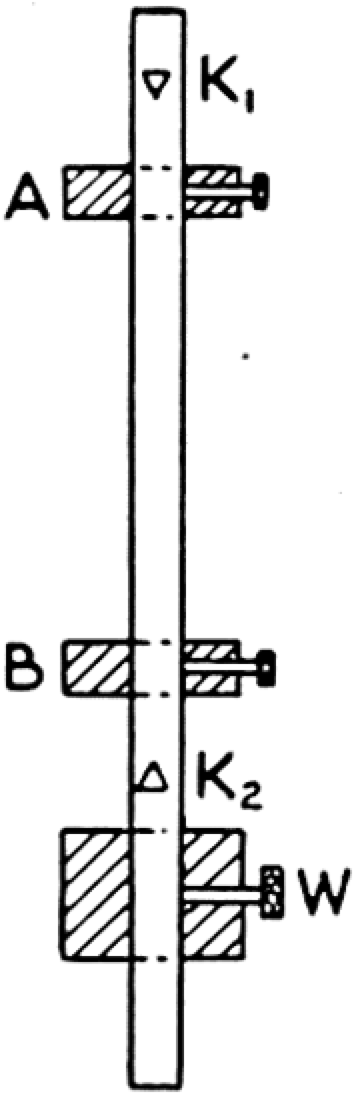


FIG. 18.

(*a*) *Absolute measurements of g.*—Various simple methods of measuring g to an accuracy of about 1 per cent. are available, such as timing the rolling of a ball down an inclined plane or the oscillations of a ball in a concave mirror whose radius of curvature is determined optically, for which the reader is referred to books on practical physics, *e.g.* Searle, *Experimental Physics*.

The ordinary laboratory method of approximating to a simple pendulum by a heavy bob attached to a thin wire or thread (**Borda's** pendulum) is capable of an accuracy of 0·1 per cent. (1 unit in g measured in C.G.S. units) but various uncertain factors, such as the stiffness of the wire and the fact that the tension, and therefore the length of the wire does not remain quite constant during the swing, make it difficult to improve the accuracy very much further. All work of the highest accuracy uses **Kater's** reversible pendulum or various improved version of it. In its original form (Fig. 18) it consisted of a rod carrying the bob W, two adjustable weights A and B, and two knife-edges K_1 and K_2, either of which can be used as a pivot. To make a determination of g, A is adjusted until the times of swing about the two knife-edges are nearly equal and then the position of B is gradually altered by means of a screw to make the periods as nearly equal as possible. If they were exactly equal then (see p. 32) the distance between K_1 and K_2 would be the length of the simple equivalent pendulum.

It is, however, unnecessary to make an adjustment to exact equality of period. Suppose that the two periods t_1 and t_2 are very nearly equal, and that the distances of K_1, K_2 from the centre of gravity are l_1, l_2, respectively, then we have (p. 32)

$$\frac{gt_1^2}{4\pi^2} = \frac{k^2 + l_1^2}{l_1} \quad \text{and} \quad \frac{gt_2^2}{4\pi^2} = \frac{k^2 + l_2^2}{l_2}$$

(where k is the radius of gyration about the centre of gravity)

so that

$$\frac{g}{4\pi^2}(t_1^2 l_1 - t_2^2 l_2) = l_1^2 - l_2^2$$

or

$$\frac{4\pi^2}{g} = \frac{t_1^2 l_1 - t_2^2 l_2}{l_1^2 - l_2^2} = \frac{1}{2}\left[\frac{t_1^2 + t_2^2}{l_1 + l_2} + \frac{t_1^2 - t_2^2}{l_1 - l_2}\right] \quad \ldots\ldots\ldots \quad (1)$$

Now $l_1 + l_2$ is the distance between the two knife-edges K_1, K_2, which can be measured extremely accurately, *e.g.* by a travelling microscope or, in extremely accurate work, by an interferometric method. Since t_1 can be made very nearly equal to t_2, the term involving $l_1 - l_2$ is only a correcting term and $l_1 - l_2$ can be determined with sufficient accuracy by balancing the pendulum about a knife-edge to determine the position of the centre of gravity. The design of the pendulum must be such that l_1 and l_2 are not nearly equal, for, if they were, the correcting term would not be small.

In order to take proper advantage of an accurate measurement of $l_1 + l_2$ a very accurate method of timing is required. The accuracy of the ordinary laboratory method of timing by stop-watch is not much better than a half-second, so that several thousand swings would have to be counted in order to get a result reliable to 1 part in 10,000, during which time the temperature of the laboratory may change, with a consequent alteration in the length of the pendulum. A much better method of timing is the *method of coincidences*. Let us suppose that the pendulum of a standard clock is made to actuate a neon lamp at a certain point of its swing. This lamp is made to illuminate a slit which is closed by a pointer on the Kater pendulum when at the centre of its swing but is open at all other times. Let the Kater pendulum be adjusted to have a period of slightly greater than 1 sec.—suppose, for example, that its period is known to be in the neighbourhood of 1·03 sec. Then, suppose the pendulum is started and observations are begun when a coincidence occurs, *i.e.* when the slit is just blocked by the pointer at the moment of flash. After 33 more swings very nearly 34 sec. will have elapsed and an approximate coincidence occurs. After 33 more swings another approximate coincidence occurs, but each time the pendulum gains slightly on the clock. If N such sets of 33 swings are observed, until the pendulum has gained one whole second on the clock, so that the next coincidence occurs in the middle of the 32nd swing, we have

$$33Nt = 34N - 1$$

so that

$$t = \frac{34}{33} - \frac{1}{33N}.$$

The method can be further refined if a scale is attached to the pointer of the pendulum and arranged to be illuminated by each flash of the lamp, thus enabling the angular distance of the pendulum from the centre of the swing to be noted at each approximate coincidence. Allowing for this we may say that

$$33Nt = 34N - 1 + x$$

where x is the fraction of a swing by which the position of the last coincidence differs from that of the first. This gives

$$t = \frac{34}{33} - \frac{1 + x}{33N}.$$

If N is 20 and x can be estimated correct to one-hundredth of a whole swing, this measurement gives us t with an accuracy of better than 1 in 600,000, which is far above the accuracy that can be obtained from the ordinary method of timing. Moreover, it is possible to carry on a continuous series of observations and any systematic change in t, *e.g.* due to drift of temperature, will be revealed at once by corresponding variations in N as time goes on.

Measurements of this order of accuracy entail corrections being made for a large number of disturbing factors, to enable full advantage to be taken of the very accurate measurements of the length $l_1 + l_2$ and of the periods. The chief corrections are as follows:

(1) *Finite arc of swing.*—It is shown in Chapter II, Art. 6 (p. 28), that a pendulum oscillation of finite but small amplitude a has a period

proportional to $1 + \frac{a^2}{16} + \ldots$, so that this correction can easily be made if a is measured.

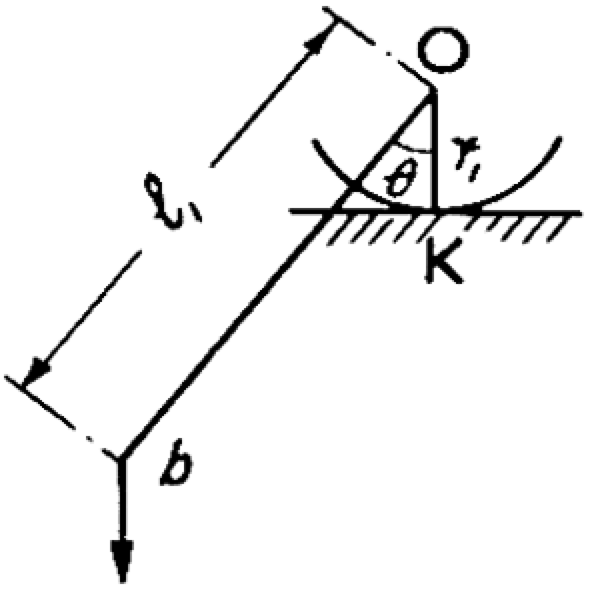


FIG. 19.

(2) *Curvature of knife-edges.*—If the knife-edges are not perfectly sharp, the pendulum swings by rolling contact of the knife-edge on its support (Fig. 19). Since two successive normals to a surface intersect at the radius of curvature, the restoring moment, if the pendulum is displaced through an angle θ, is $Mg\,(l + r)\sin\theta$, where r is the radius of curvature, but the instantaneous axis of rotation (p. 33) still passes through the bottom of the knife-edge K, so that the moment of inertia is still to be taken about this axis. We have therefore

$$M\,(k^2 + l^2)\frac{d^2\theta}{dt^2} + Mg\,(l + r)\sin\theta = 0 \quad \ldots\ldots\ldots\ldots \quad (2)$$

as the equation of motion, in place of equation (29), Chapter II, and the periods t_1, t_2 are therefore given by

$$\frac{g}{4\pi^2}t_1^2 = \frac{k^2 + l_1^2}{l_1 + r_1} = \frac{k^2 + l_1^2}{l_1}\left(1 - \frac{r_1}{l_1}\right) \text{ approx.} \quad \ldots\ldots \quad (3)$$

$$\frac{g}{4\pi^2}t_2^2 = \frac{k^2 + l_2^2}{l_2 + r_2} = \frac{k^2 + l_2^2}{l_2}\left(1 - \frac{r_2}{l_2}\right) \quad \ldots\ldots\ldots\ldots \quad (4)$$

so that

$$\frac{g}{4\pi^2}(t_1^2 l_1 - t_2^2 l_2) = l_1^2 - l_2^2 - \frac{k^2 + l_1^2}{l_1}r_1 + \frac{k^2 + l_2^2}{l_2}r_2.$$

Now if t_1 and t_2 are nearly equal, so are $(k^2 + l_1^2)/l_1$ and $(k^2 + l_2^2)/l_2$ and we may replace them both by their mean l in the terms involving r_1 and r_2. Thus we have

$$\frac{g}{4\pi^2}.\left[\frac{t_1^2 l_1 - t_2^2 l_2}{l_1 - l_2}\right] = l_1 + l_2 + \frac{l}{l_1 - l_2}(r_2 - r_1) \quad \ldots\ldots \quad (5)$$

If it is possible to interchange the knife-edges without upsetting the geometry, we shall have, if the two periods are now t_1', t_2',

$$\frac{g}{4\pi^2}\left[\frac{t_1'^2 l_1 - t_2'^2 l_2}{l_1 - l_2}\right] = l_1 + l_2 + \frac{l}{l_1 - l_2}(r_1 - r_2) \quad \ldots\ldots \quad (6)$$

so that, by addition, the correcting terms involving r_1 and r_2, which are difficult to measure directly, disappear. This rather tiresome correction can be avoided by having plane bearings on the pendulum and a knife-edge on the support, when r_1 and r_2 are equal, so that the correcting term disappears.

(3) *Yielding of the support.*—It might be thought a simple matter to devise a support that will not yield appreciably when so light a body as a pendulum, weighing perhaps 1 kg., is swung upon it, but we must remember that we have methods available for measuring lengths and periods to an accuracy of one part in many thousands. Now a mass of, say, 100 kg. hung on any ordinary support would cause quite an appreciable deflection, so that it is clear that the effect of the pendulum in deflecting the support cannot be neglected in work of high accuracy. The chief disturbing factor is the forced oscillation of the support in a *horizontal* direction. The corresponding vertical oscillation is very much less important. In Fig. 20, let O be the point of support and G the centre of gravity, so that $OG = l_1$. Suppose the horizontal yield of the support, if unit force is applied, is a. The acceleration along GG' is $l_1 \dfrac{d^2\theta}{dt^2}$, and thus the force along GG' is $ml_1\dfrac{d^2\theta}{dt^2}$, which is equal to the force on the support.

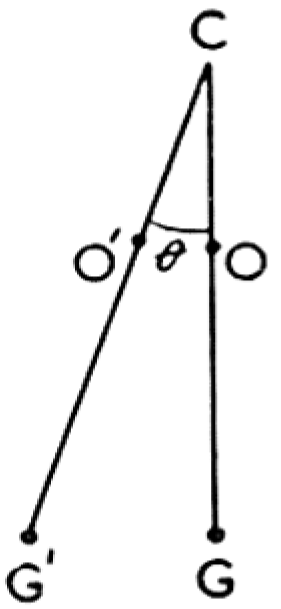


FIG. 20.

But

$$\frac{d^2\theta}{dt^2} = \frac{gl_1\theta}{k^2 + l_1^2}$$

so that the force on the support is

$$\frac{mgl_1^2\theta}{k^2 + l_1^2} = \frac{mgl_1\theta}{l_1 + l_2}$$

since $k^2 = l_1l_2$ if the periods for both knife-edges are equal (p. 55). The yield OO' is thus related to the angular deflection θ by

$$OO' = amg\frac{l_1\theta}{l_1 + l_2}$$

so that the centre of oscillation is effectively raised from O to C, where

$$\frac{OO'}{\theta} = OC = \delta_1 = \frac{amgl_1}{l_1 + l_2}.$$

The period is therefore given by

$$\frac{g}{4\pi^2}t_1^2 = \frac{(l_1 + \delta_1)^2 + k^2}{(l_1 + \delta_1)} = l_1 + \delta_1 + \frac{k^2}{l_1 + \delta_1}.$$

For oscillation about the other knife-edge

$$\frac{g}{4\pi^2}t_2^2 = l_2 + \delta_2 + \frac{k^2}{l_2 + \delta_2}$$

whence

$$\frac{g}{4\pi^2}\left[\frac{t_1^2 l_1 - t_2^2 l_2}{l_1 - l_2}\right] = \frac{1}{l_1 - l_2}\left[l_1^2 + \delta_1 l_1 + \frac{k^2 l_1}{l_1 + \delta_1} - l_2^2 - \delta_2 l_2 - \frac{k^2 l_2}{l_2 + \delta_2}\right]$$

$$\simeq l_1 + l_2 + \frac{1}{l_1 - l_2}(\delta_1 l_1 - \delta_2 l_2)$$

$$= l_1 + l_2 + mga \quad \ldots\ldots\ldots\ldots \quad (7)$$

using the above expression for δ_1, and the corresponding one for δ_2.

Thus the effective length of the equivalent simple pendulum is increased by mga, which is the deflection of the support produced by a horizontal force equal to the weight of the pendulum. a can be determined by applying static loads to the support and observing its deflection by means of a microscope. This correction can be reduced considerably by having two pendulums swinging in antiphase on the same support, this method being due to **Meinesz**, but this involves adjusting the pendulums until their periods are only very slightly different from one another. A method of eliminating this correction is described below (p. 60).

(4) *Effects of the air.*—There are three distinct ways in which the presence of air may affect the period. First, the buoyancy of the air has to be allowed for. This is quite simple. If m is the mass of the pendulum and m' that of the air it displaces, then the restoring couple is diminished from $mgl \sin\theta$ to $(ml - m'\delta)\, g \sin\theta$, where δ is the distance of the centre of gravity of the displaced air from the axis of rotation, which can be calculated with sufficient accuracy from the dimensions of the pendulum. (Note that δ is not necessarily equal to l unless the density of the pendulum is uniform.) Secondly, we have to allow for the damping due to the viscosity of the air. On the assumption that the damping couple can be represented by a term proportional to the angular velocity of the pendulum, one can easily compute a relation between the ratio of amplitudes of successive swings and the increase of the period. The same thing can be done, though less simply, if the damping couple is proportional to the square of the angular velocity, and there is reason to believe that the true law lies between the two. In either case, the effect on the period is completely negligible for a pendulum moving in air (see Appendix, p. 287).

The third correction arises from the fact that a pendulum moving through the air "carries air with it". For a body being accelerated through a medium in a straight line work has to be done in imparting kinetic energy to the medium in addition to that supplied to the body, and it may be supposed that the velocity of any portion of the medium is proportional to the velocity of the body. If so, the total kinetic energy of the medium is proportional to the square of the velocity of the body, and the effect of the medium can be represented by introducing the idea of "virtual mass" (see p. 164), the additional mass that has to be added to the body to represent the effect of the medium. A similar conception of "virtual moment of inertia" applies to a body *rotating* in a medium. Only for a few simple shapes of body, *e.g.* ellipsoids, can the effect of the medium be evaluated accurately (see Lamb's *Hydrodynamics*), and here it is found that the virtual moment of inertia is proportional to the density of the medium, as one would expect. The true equation of motion, neglecting the damping which has already been considered, but taking account both of the buoyancy and the virtual mass, is therefore

$$m\left(k^2 + l^2 + \frac{m'}{m}K\right)\frac{d^2\theta}{dt^2} + g\,(ml - m'\delta)\sin\theta = 0$$

where K is a constant that depends on the size and shape of the pendulum, the order of magnitude of which can be found by considering an ellipsoid of dimensions similar to those of the pendulum. K can also be evaluated by working with two pendulums of different masses, but of the same size and shape. However, **Repsold** constructed a pendulum of Kater's type, but symmetrical in geometrical form about its mid-point, for which the corrections due to the medium, apart from the damping effect, can be avoided. For we have

$$\frac{g}{4\pi^2}t_1^2 = \frac{k^2 + l_1^2 + K_1}{l_1\left(1 - \frac{m'}{m}\frac{\delta_1}{l_1}\right)}$$

$$= \frac{k^2 + l_1^2}{l_1} + \frac{k^2 + l_1^2}{l_1}\frac{m'}{m}\frac{\delta_1}{l_1} + \frac{K_1}{l_1} + \ldots$$

neglecting powers and products of small quantities. Also

$$\frac{g}{4\pi^2}t_2^2 = \frac{k^2 + l_2^2}{l_2} + \frac{k^2 + l_2^2}{l_2}\frac{m'}{m}\frac{\delta_2}{l_2} + \frac{K_2}{l_2},$$

from which we have

$$\frac{g}{4\pi^2}\frac{t_1^2 l_1 - t_2^2 l_2}{l_1 - l_2} = l_1 - l_2 + \frac{m'l}{m}\frac{\delta_1 - \delta_2}{l_1 - l_2} + \frac{K_1 - K_2}{l_1 - l_2}$$

where l has been written in the small term for $\frac{k^2 + l_1^2}{l_1}$ or $\frac{k^2 + l_2^2}{l_2}$, which are approximately equal. If now the geometry of the pendulum is the same from whichever knife-edge it is suspended, we have $K_1 = K_2$ and $\delta_1 = \delta_2$. For the validity of all the corrections it is obvious that l_1 and l_2 must *not* be nearly equal otherwise the correcting terms would not be small. In Kater's original pendulum this was secured by making one bob much larger than the other, and the same thing can be done with Repsold's pendulum by making one bob solid and the other hollow.

It is possible to develop this method of correction one step further, as was done by **Deffrages,** to eliminate the troublesome correction due to the yield of support. He used two pendulums of Repsold type, of equal weights but of different lengths, designed so that the ratio l_1/l_2 was the same for both. We have already shown how the use of a Repsold type of pendulum eliminates the necessity of correcting for the buoyancy and virtual mass of the air. We also suppose the same pair of knife-edges used for both pendulums, and we have shown above (p. 58) that the effect of a yielding support on the length of the equivalent simple pendulum is the same for both pendulums, since it depends only on the weight of the pendulum. Then, if T_I and T_{II} represent the periods of the two pendulums I and II of different lengths, supposing for simplicity that each has been adjusted to exact equality of period about either knife-edge, we have, from equation (5)

$$\frac{g}{4\pi^2} T_I{}^2 = (l_1 + l_2)_I + \delta + \left(\frac{l_1 + l_2}{l_1 - l_2}\right)_I (r_2 - r_1),$$

$$\frac{g}{4\pi^2} T_{II}{}^2 = (l_1 + l_2)_{II} + \delta + \left(\frac{l_1 + l_2}{l_1 - l_2}\right)_{II} (r_2 - r_1),$$

the correction δ being that for yield of support considered above which will disappear on subtracting these two equations, as will the correction for the curvatures of the knife-edges, if we use the same pair for both pendulums, and make the ratio l_1/l_2 the same for both.

We have gone into the question of applying these corrections in some detail, because it is a typical illustration of the new problems that arise in physics at each stage of improving the accuracy of an absolute measurement. Very typical also is the use of a *differential* method, like the use of the reversible pendulum, in simplifying, reducing, or eliminating, the corrections that have to be made. For many purposes, an absolute value of g is not needed, one wants instead the *variation* of g from point to point on the earth's surface, and various instruments have been devised for rapid measurements of this variation. If what is wanted is the *ratio* of g at two places, many of the corrections can be dispensed with, and the times of swing of robustly made compound pendulums compared at the two places. To reduce errors due to thermal expansion they are made of the alloy invar, which has a very small coefficient of expansion, and air currents are eliminated by using an evacuated box.

(*b*) *Measurements of the differences in g between two stations.*—A simple instrument of this type is the quartz gravity balance. It consists essentially of a fine quartz fibre supporting a wire, the wire carrying a small weight so that its centre of gravity is on one side of the fibre (Fig. 21). The fibre itself is kept in tension by means of a spring, and can be twisted by means of a torsion head until the wire is horizontal. Any further twist would cause the wire to rotate completely. The approach of this instability can be detected easily, for it will occur when the torsional couple due to the fibre just balances that due to the pull of the earth when the wire is horizontal and, when this state of affairs is approaching, the net couple acting on the wire only varies slowly as its inclination changes, so that the torsional oscillations of the wire about its equilibrium position become *of longer and longer period* as instability approaches. Thus, if temperature corrections are made for the expansion of the wire and the change in rigidity of the fibre, the difference in g between two stations is proportional to the difference in torsional couples required to produce instability. The instrument can be calibrated by observations at two stations where g is accurately known. Instruments on a similar principle, but using a sensitive spring balance in place of the torsion fibre, are also much used in gravity surveys.

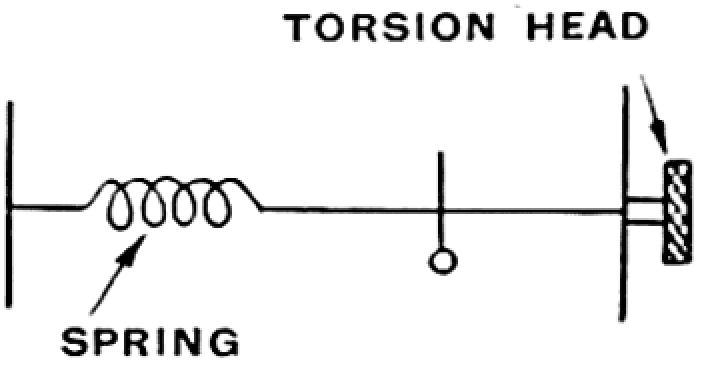


FIG. 21.

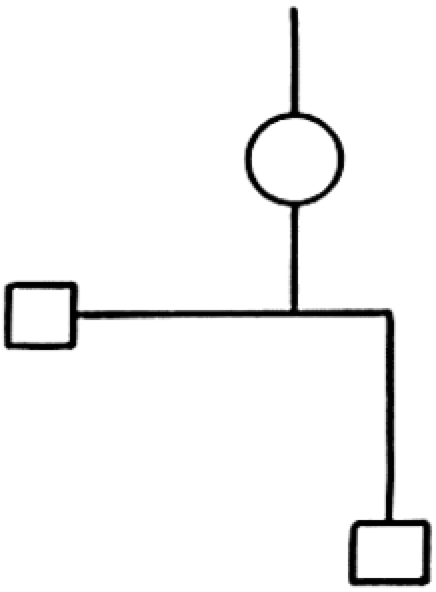


FIG. 22.

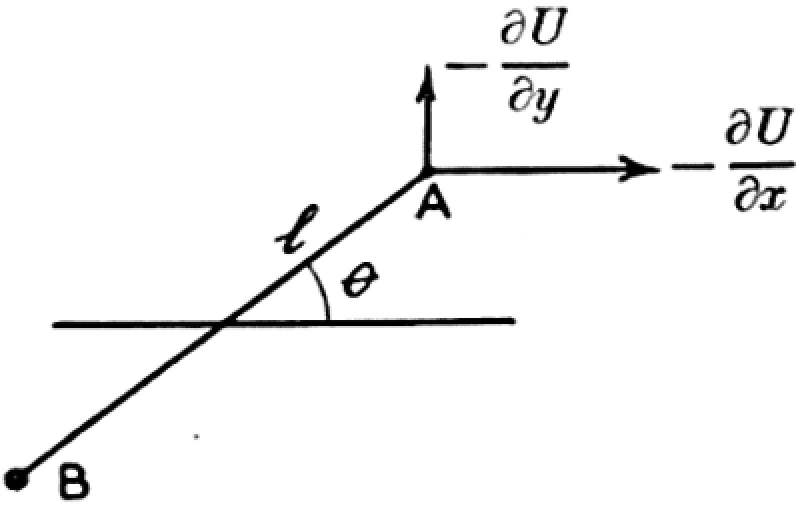

$-\frac{\partial U}{\partial y}$
$-\frac{\partial U}{\partial x}$
A
l
θ
B

FIG. 23.

(*c*) *Direct measurements of the gradient of g. The Eötvös balance.*—The torsion balance due to **Eötvös** consists basically of two masses at different levels, connected by a bar and suspension wire, the whole being suspended on a torsion wire furnished with a mirror (Fig. 22). First, consider a simpler system, in which both masses are suspended at the same level, and suppose first that the mass of the connecting rod is negligible (Fig. 23). Then, if U is the gravitational potential, and the beam is inclined at an angle θ to the coordinate axes, the force on mass A in the x-direction is $-m\frac{\partial U}{\partial x}$, taken at the position of A which is

$$x = l\cos\theta \qquad y = l\sin\theta.$$

The *moment* of this force about the origin is

$$-m\left(\frac{\partial U}{\partial x}\right)_A l\sin\theta.$$

The *moment* of the corresponding force on mass B is

$$m\left(\frac{\partial U}{\partial x}\right)_B l\sin\theta.$$

The total moment due to the *variation* of $\partial U/\partial x$ as we go from A to B is therefore—

$$-ml\sin\theta\left[\left(\frac{\partial U}{\partial x}\right)_A - \left(\frac{\partial U}{\partial x}\right)_B\right].$$

It is to be especially noticed that $\partial U/\partial x$, which one might call the "x component of g", produces no observable effect, but merely produces an additional *force* on the system and draws it *aside* slightly, as in the mountain experiments, but does not produce any *torsion* in the wire. In no circumstances, therefore, can this instrument measure the first derivatives of U, or g itself, directly. The torsion arises from the *variation* of these first derivatives with *distance*. The moment due to the variation of $\partial U/\partial y$ as we go from A to B is similarly

$$+ml\cos\theta\left[\left(\frac{\partial U}{\partial y}\right)_A - \left(\frac{\partial U}{\partial y}\right)_B\right].$$

The total couple C on the wire due to these variations is, if we suppose the second derivatives of U constant,

$$C = -ml\sin\theta\left[\frac{\partial^2 U}{\partial x^2}2l\cos\theta + \frac{\partial^2 U}{\partial x.\partial y}2l\sin\theta\right]$$

$$+ml\cos\theta\left[\frac{\partial^2 U}{\partial y^2}2l\sin\theta + \frac{\partial^2 U}{\partial x.\partial y}2l\cos\theta\right]$$

by expanding $\partial U/\partial x$ and $\partial U/\partial y$ by Taylor's theorem, since

$$x_A - x_B = 2l\cos\theta \quad \text{and} \quad y_A - y_B = 2l\sin\theta$$

or $$C = -ml^2\sin 2\theta\left(\frac{\partial^2 U}{\partial x^2} - \frac{\partial^2 U}{\partial y^2}\right) + 2ml^2\cos 2\theta\left(\frac{\partial^2 U}{\partial x.\partial y}\right) \quad \ldots (8)$$

If we replace the concentrated masses by a bar, each element of mass on the bar will contribute to the torque in proportion to the square of its distance from the suspension, so that in expression (8) we may replace $2ml^2$ by I, the moment of inertia of the suspended system. If, further, we suppose that mass B is suspended at a distance h below A, then

$$\left(\frac{\partial U}{\partial y}\right)_A - \left(\frac{\partial U}{\partial y}\right)_B = \frac{\partial^2 U}{\partial y^2}\, 2l \sin\theta + \frac{\partial^2 U}{\partial x . \partial y}\, 2l \cos\theta + h\frac{\partial^2 U}{\partial y . \partial z}$$

while $\left(\frac{\partial U}{\partial x}\right)_A - \left(\frac{\partial U}{\partial y}\right)_B$ contains a similar extra term $h\frac{\partial^2 U}{\partial x . \partial z}$, and the couple contains the terms

$$mh \cos\theta \frac{\partial^2 U}{\partial y . \partial z} - mh \sin\theta \frac{\partial^2 U}{\partial x . \partial z}$$

in addition to the two given in expression (8).

These four contributions to the couple involving

$$\left(\frac{\partial^2 U}{\partial x^2} - \frac{\partial^2 U}{\partial y^2}\right), \quad \frac{\partial^2 U}{\partial x . \partial y}, \quad \frac{\partial^2 U}{\partial x . \partial z}, \quad \text{and} \quad \frac{\partial^2 U}{\partial y . \partial z}$$

can be distinguished from one another because they vary in different ways with θ. To determine them all, it is necessary to observe the orientation of the bar with the torsion head set at *five* different azimuths, thus giving us four equations between the four unknown quantities. *Five* azimuths are needed, not four, because the inclination that the bar would make with the pointer on the torsion head if there were no local variation of gravity (the "zero reading" of the instrument) is not known, so that the setting of the torsion head, and therefore the applied torsional couple, has to be measured from an arbitrary zero, and a fifth equation is needed to determine the true zero.

We require to know I, the moment of inertia of the suspended system, and the torsional constant of the suspension. This latter quantity can be determined by placing an accurately-turned lead sphere to the right and left of the beam at a measured distance from it, thus producing a known change in the local gravitational field. It will be observed that there are two pieces of information missing when we have made a survey with the Eötvös instrument, namely, $\partial^2 U/\partial z^2$ and either $\partial^2 U/\partial x^2$ or $\partial^2 U/\partial y^2$ separately. There is, however, the equation

$$\frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} + \frac{\partial^2 U}{\partial z^2} = 0 \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (9)$$

which applies to any region in which matter is not present, because the gravitational field is of inverse-square type. For a proof that equation (9) is equivalent to the statement that the potential U refers to an inverse-square field see Jeans, *Electricity and Magnetism.* There is thus only *one* of the six second derivatives of U left undetermined by a survey with this instrument, though as stated already, it can give no information at all about the value of g itself.

(*d*) *Measurement of g at sea.*—The earliest attempts were based on the comparison of the readings of a mercury barometer, whose reading varies with g, with those of either an aneroid barometer or a hypsometer (that is an instrument that measures the atmospheric pressure by the boiling point of water). To get a result accurate to a tenth of a unit in g, the reading of the mercury barometer has to be accurate to within a tenth of a millimetre. This is easy enough under laboratory conditions but becomes difficult even under conditions which sailors would describe as "glassy calm". Modern gravity surveys rely on a pendulum method, in which the accelerations caused by the rolling and pitching of the ship can be allowed for to a first approximation by subtracting from the motion of the pendulum used in the measurement that of one that is initially at rest. Since the correction so made is only valid to the first order of small quantities, the second order corrections are kept as small as possible by inclination of the box containing the pendulums according to the motion of the ship. It is possible to do this manually, but it is usually done by means of a servo-mechanism. If this is properly designed, the disturbances to the pendulums can be greatly reduced.

Another very recent method consists of measuring the change in frequency of the transverse vibrations of a wire stretched by a weight. Variations of g do not affect the *density* of the wire, which depends on its *mass*, but do affect the pull of the earth on the weight, and therefore the *tension* in the wire. Variations of frequency are detected by comparing the vibrations of the wire, by a beat-frequency method, with those of a quartz crystal oscillator, whose frequencies are independent of g. A change of frequency of a small fraction of a vibration per second can be detected by such means even for a frequency of several thousands per second, so that such an instrument is, in principle, capable of great accuracy. This method is also applicable to measurements on land, and is likely to be of great value, as it can be used in awkward places such as the bottoms of boreholes.

6. Applications of the gravity survey

Let us suppose that one is trying to locate a coal-measure, or a deposit of oil, by means of a gravity survey, relying on the difference in density of the coal or oil and the overlying rocks. It is not to be expected that one can obtain *precise* indications of the extent and thickness of the deposit from gravity data alone. We know, for example, that no observations of gravity *outside* a spherical mass could distinguish between a sphere of uniform density and a hollow sphere of the same mass, or a sphere with a heavy core (p. 49).

By means of the Eötvös balance, or an instrument of torsion or spring-balance type, one can map one's district, *e.g.* in "contours" of constant g, or by drawing "gravitational lines of force" analogous to electric lines of force. We must obviously eliminate all variations of g due to *known* causes before being in a position even to speculate on what distribution of matter below the surface could produce the observed effects. The various corrections that are necessary are as follows:—

(*a*) *Topographical correction.*—It is customary to reduce data to what would be expected if the land immediately surrounding the station were a horizontal plane. This correction is always positive. For let Fig. 24 represent a section through the terrain surrounding station S. We require to reduce our observations of g to what we would expect if the actual terrain were replaced by the horizontal plane ASB. The centre of mass of the region $ASRQP$ is above AB, and therefore attraction at S has an *upward* component which diminishes g, so that g would be *increased* if this region were removed, while the centre of mass of the region BST is below AB, so that its attraction at S would have a downward component, and *adding* this region would *increase* g.

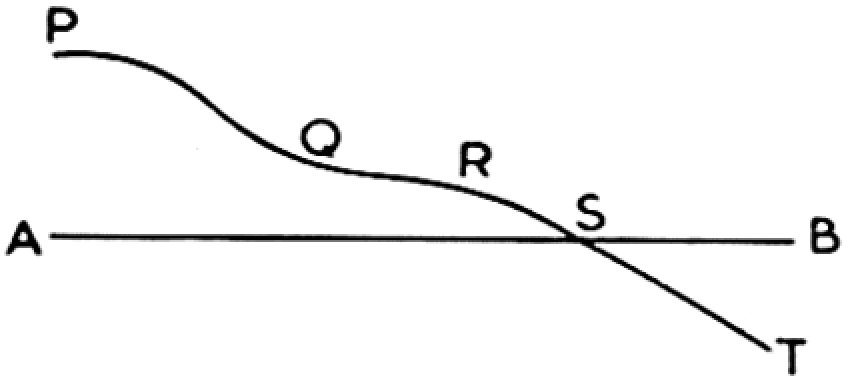


FIG. 24.

(*b*) *Corrections for altitude and latitude.*—The correction for altitude simply involves taking account of the fact that the observer is further from the centre of the earth, which, applying the inverse-square law, suggests the simple correction $g_h = g_o (1 - 2h/R)$. To correct for latitude we have to allow for the centrifugal effect of the earth's rotation which gives a centrifugal acceleration $\omega^2 R \cos \lambda$, where R is the radius of the earth and λ the latitude. Resolving this acceleration vertically, we get a correction to g proportional to $\cos^2 \lambda$. The ellipticity of the earth necessitates a further correction, which is shown in text-books on potential theory to be proportional to $\cos^2 \lambda$ also. The latitude correction is not necessary unless a survey is being made on a large scale, *e.g.* to examine conditions near a mountain range or the edge of a continent.

(*c*) *Correction for elevated masses.*—The altitude correction $1 - 2h/R$ is only strictly valid for an observer in a balloon, as it assumes that the observer is at a height h above a spherical earth. If he gains this height by climbing a mountain, the local effect of the mountain ought to be allowed for. At first one might think that this could be done in a manner similar to the topographical correction. Bouguer, indeed, suggested the modified correction

$$g_h = \left(1 - \frac{2h}{R} + \frac{3}{2}\frac{h}{R}\frac{d}{D}\right)$$

where d is the density of the mountain and D the mean density of the earth. This correction is based on the same idea that inspired the original plumb-line experiments for measuring G or D, namely, that the attraction of the mountain of density d could simply be added to that of a spherical earth of mean density D. It is, however, now known that the net effect of a mountain is often much less than Bouguer thought, and it is this fact that prevented the plumb-line experiments from giving more than "order of magnitude" determinations of G. The correction for the presence of a mountain being rather uncertain, it is difficult to use gravity surveys in a mountainous district for the purpose of locating mineral deposits. Bouguer's correction is an upper limit to the possible effect of a mountain, and one can get a lower limit by neglecting its attraction altogether. Such an assumption is by no means so foolish as it sounds, as we shall see in the next article.

7. Joly's theory of the rise and fall of the continents

Joly propounded a very remarkable theory to account for the undoubted geological fact that sedimentary rocks such as chalk and sandstone, which can be formed only under the sea, often exists at quite high altitudes. He supposed that the rocks beneath the surface crust of the earth are heated, by the breakdown of radioactive elements, until they melt, whereupon they expand and their density decreases and the overlying continental masses sink into them like icebergs into water. The rock has expanded on liquefying, and the sinking of the continents further reduces the volume available and causes the liquid rock to "well up" round them, whereupon it loses its heat to the sea by conduction and gradually solidifies, its density increases, and the continents rise again, the cycle repeating itself when a fresh store of radioactive heat has built up under them. During this up and down motion the continents crumple, leading to the formation of mountain ranges. Geologically it appears that there were quite a number of of epochs of "mountain-building" after the earth's crust solidified.

If this theory is correct, it means that a particular range of mountains may at present be in any condition between being supported like icebergs on an effectively liquid substratum and simply resting on an underlying bed of solid rock. If we assume the former state of affairs, we could "smooth off" the earth's surface by replacing the mountain by an equal mass of the slightly denser rock in which the mountain floats, and since the centre of gravity of the mountain, *including the portion of it submerged beneath the rock in which it floats,* must be fairly near that of the mass of rock it displaces, one would expect the effect of the mountain on the local variation of gravity to be small. If we assume that the mountain is resting on an absolutely rigid layer of rock, then Bouguer's assumption would be correct.

Gravitational data near mountains are sometimes described as indicating "compensation" at a certain depth, that is to say, are consistent with what would be expected if there were hydrostatic equilibrium at and below that depth, but not right up to the surface of the earth as the "iceberg" assumption would require, nor at very great depths only, as Bouguer's assumption implies. The conception of a "depth of compensation" is a convenient way of describing a state of affairs between these two extremes, zero depth of compensation implying the "iceberg" assumption, infinite depth of compensation, Bouguer's assumption. From the evidence of earthquake propagation it is improbable that the whole substratum is liquid, but there is no reason why some of it should not be. There is a strong defect of gravity over the "earthquake zone" at the eastern and southern edges of Asia, which may indicate that the continental mass in this region is in the process of sinking, perhaps due to Joly's mechanism, and the adjustment in the geological strata that accompanies the sinking leads to frequent earthquakes in this region, particularly in Japan.

8. Einstein's theory of gravitation

Although we have assumed Newton's law of gravitation throughout this chapter, all the numerical consequences of it so far deduced remain true if Einstein's law is substituted. It is only for very rapidly moving bodies that any observable difference arises. The chief difference between Einstein's and Newton's theories is one of how the phenomena of gravitation is to be most consistently described. Newton, following Galileo, regarded gravitation as a *force* occurring in the neighbourhood of matter, while Einstein considers that the presence of an attractive mass fundamentally affects the *geometrical properties of the space* in its neighbourhood. The following are the three phenomena which are usually held to give decisive evidence for Einstein's theory.

(*a*) *The advance of the perihelion of Mercury.*—It has long been known that the orbit of Mercury does not simply repeat itself year after year, as it should on the basis of an inverse-square law of attraction (p. 26), but is continually changing its position, so that the actual path of the planet is of "rosette" shape. Part of this effect can be accounted for by the attraction of other planets but the expected effect due to this can be calculated precisely and is definitely too small. At one time the residuary effect was ascribed to an intra-Mercurial planet, "Vulcan", but all attempts to find such a planet have failed. The Einstein theory does predict an effect of the correct size, which arises from the high orbital velocity of Mercury, which results in a noticeable departure from Newton's inverse-square law. The increase in the apparent mass of a particle at high velocities, predicted by Einstein, is familiar to all nuclear physicists.

(*b*) *The deflection of a light-ray passing near the sun.*—This effect has been detected by photographing certain stars near the sun during an eclipse, and the same stars again when the sun has moved to another part of the heavens. On Einstein's theory, a distortion of the geometry of space near the sun implies that the path of a light-ray will be curved in its neighbourhood, but it must be emphasised that almost any *corpuscular* theory of light would also imply a deflection of a light-ray by a heavy body. For we pointed out above (p. 50) that the orbit of a small body about a heavy one is *independent of the mass* of the small body, so that the deflection of a light-corpuscle would be the same, on Newton's theory of light, whatever its mass. It follows that the comparison between Newton's and Einstein's theories is here only a numerical one, the comparison being possible because Einstein predicts twice as large an effect as does Newton's corpuscular theory of light, while Maxwell's electromagnetic theory of light combined with Newtonian gravitation predicts no effect at all. Not all of the measurements that have been made during eclipses point decisively to Einstein's figure, but, making allowance for the fact that the conditions for some observations were decidedly more favourable than for others, the balance of evidence does point towards Einstein's figure. Einstein's theory, moreover, enables us to retain Maxwell's electromagnetic theory of light, which is supported by much evidence.

(*c*) *The increase in wave-length of the light from a star.* (*The Einstein shift.*)—By combining the theory of relativity and quantum theory we can calculate precisely the mass of a light-quantum of given frequency. The quantum is attracted by the gravitational field of the star, and loses some energy in travelling through it. According to Planck's law, this effect leads to a slight drop in the frequency. We have, by Planck's law

$$E_s = h\nu_s$$
$$E_o = h\nu_o$$

where the suffix s refers to the surface of the star, and the suffix o to a point at a great distance. By Einstein's relation between the mass of a body and its total energy, for which the reader is referred to textbooks on relativity

$$E = mc^2.$$

The work done on a body of mass m, in removing it from the surface of a star of radius R and mass M to a great distance is $\frac{GMm}{R}$ (p. 47). Applying this to a light-quantum, we have

$$\frac{GMm}{R} = \frac{GME_s}{Rc^2} = E_s - E_o.$$

We have therefore

$$\frac{E_s - E_o}{E_s} = \frac{\nu_s - \nu_o}{\nu_s} = \frac{GM}{Rc^2},$$

the percentage shift in frequency being thus independent both of h and the frequency. This result is not very easy to check, both because of the numerical smallness of the effect, and also because it requires knowledge of both M and R for the same star. M can only be measured directly for double stars, and R only for stars that are not too far away. Numerical verification of this formula has been possible for the sun and for the *companion of Sirius*. The latter star is a "white dwarf" of high density, so that M/R is comparatively large, M can be inferred from the motion of Sirius itself (p. 50), while, since Sirius is one of the nearest stars, it is possible to measure R optically by interference methods.

Thus the position can be summed up by saying that Newton's theory is an extremely close approximation, valid for all ordinary purposes, but that the balance of experimental evidence for the three phenomena mentioned is decidedly in favour of Einstein's theory. Attempts have been made to develop Einstein's theory to obtain a picture of the large-scale geometry of the universe and to explain the motion of the galaxies, but so many possibilities are open that this field of enquiry, known as "cosmology", must be regarded as very speculative indeed. There is not nearly enough evidence to decide between current theories, and the collection of more evidence is likely to prove slow and difficult.

BIBLIOGRAPHY FOR CHAPTER III

A. S. Eddington, *Space, Time, and Gravitation.* C.U.P. (1920).

A. S. Eddington, *Mathematical Theory of Relativity.* C.U.P. (1924).

J. H. Jeans, *Mathematical Theory of Electricity and Magnetism.* C.U.P. (1927).

H. Jeffreys and B. S. Jeffreys, *Methods of Mathematical Physics.* 2nd Edn. C.U.P. (1950).

H. Jeffreys, *The Earth.* 2nd Edn. C.U.P. (1929).

Poynting and Gray, *Phil. Trans. A* (1899), **192,** p. 245.

Examples on Chapter III

1. Assuming that the value of g at the surface of the earth at the equator is 980 cm./sec.^{2}, calculate the value in an aeroplane flying at 600 km./hr at a height of 8 km. above the equator: (*a*) from east to west; (*b*) from west to east.

Radius of earth = 6,500 km. (Hint.—g at the equator depends on both centrifugal effect and gravitational attraction. Separate these two before taking account of the effect of height.)

2. Show how to determine the mass of the sun, assuming that the earth's orbit is a circle of radius $1{\cdot}5 \times 10^8$ km. and that the gravitational constant is $6{\cdot}7 \times 10^{-8}$ C.G.S. units. Determine also the mass of the earth given that the mean distance of the moon is 4×10^5 km. Explain clearly why knowledge of the mass of the moon is not needed for the calculation.

3. Determine the height to which an observer must rise in a balloon so that the seconds pendulum is shortened by 1 cm. Take g at the surface of the earth as 980 cm./sec.^{2} and radius of earth as 6,500 km. (A "seconds pendulum" has a complete to-and-fro period of 2 sec.)

4. Explain the method of timing a pendulum by coincidences, pointing out its superiority to the ordinary method of timing. What factors limit the accuracy obtainable?

5. Explain clearly the various factors that determine the deflection of a plumb-line in the neighbourhood of a mountain.

6. Assuming that a method of timing a pendulum to an accuracy of one part in 100 thousand were available, describe in outline the design of pendulum you would use for a measurement of g to this accuracy and the corrections you would apply. Where applicable, indicate briefly your reasons for the choice of special materials.

CHAPTER IV

THE DEFORMATION OF SOLIDS. ELASTICITY AND PLASTICITY

1. Possible modes of deformation of a solid

The usual definition of a solid body is one that "preserves its size and shape", in contrast with liquids and gases, which have to be contained in some kind of vessel. More precisely, this means that forces have to be applied to a solid in order to change its size and shape, and that, if these forces are small enough, the solid returns to its original geometrical shape when the forces are removed. We are thus led to a study of the relationships between the deformation of a solid and the forces applied to it. We shall be mainly concerned with *isotropic* solids, that is to say solids whose properties are the same in all directions. This is not generally true of bodies with a crystalline structure. However, many of the solids that one meets with in practice either have no definite crystalline structure, *e.g.* glass, or else are composed of a multitude of very small crystals orientated in all sorts of directions, as occurs in metals in the state in which the engineer uses them. For the present we shall discuss the behaviour of isotropic solids, but we must always bear in mind that caution may be necessary in applying the resulting theory to metals. For a detailed discussion of the deformation of crystalline bodies the reader is referred to *Elasticity* by A. E. H. Love.

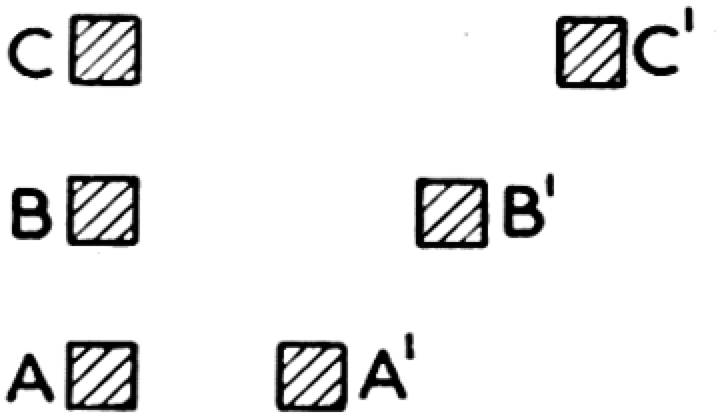


FIG. 25.

We consider first the deformation which is the same at all points of a body—that is to say, all small cubical elements of the body are deformed in the same way, so that the unstrained solid may be supposed to be built up of small equal cubes, while, in the strained state, these may have moved relative to one another, or may all have altered in size and shape in the same way. This state of affairs is called one of *uniform strain.* Now it is evident from Fig. 25 that strain can only be uniform if *points that were in a straight line in the undeformed body remain in a straight line in the deformed one.* For, let A, B, and C be three elements that were originally in the same straight line, such that $AB = BC$, and let A', B', and C' be their new positions. For the strain to be uniform $A'B'$ must obviously be equal to $B'C'$, and also $A'B'C'$ must be a straight line, otherwise the displacement of C relative to B could not be equal to that of B relative to A. From coordinate geometry we know that there are only two types of transformation of coordinates in which straight lines become straight lines. Since straight lines are represented by equations of the first degree, they can only remain straight lines if the equations governing the transformation of coordinates are of the first degree also, so that the most general type of transformation is of the following type:—

$$x' = ax + by + cz, \text{ etc.}$$

Geometrically, any such transformation can be arrived at by a superposition of operations of two distinct types:

(*a*) A change in the *scale* of the coordinates. The corresponding type of strain is called a *dilatation*, which results in an originally cubical element becoming cuboidal. (Of course a dilatation can be *uniform from point to point in a body* without being the *same in all three directions*.)

(*b*) A change in the *inclinations* of the coordinate axes. The corresponding type of strain is called a *shear*. One may regard this either as a relative *sliding* of the cubical elements over one another, in the manner of a pack of cards pushed sideways, or as the *deformation* of originally cubical elements into *rhomboids* without *change of size*.

It therefore follows that any uniform strain can be decomposed into three *dilatations*, one for each of the three coordinate axes, and three *shears*, one corresponding to each of the three pairs of coordinate axes. The three dilatations are usually represented by the symbols, e_{xx}, e_{yy}, e_{zz}, where, for example, e_{xx} represents the *fractional elongation* in the direction of the x-axis. Similarly, e_{xy} or e_{yx}, is defined as the *change in inclination* (measured in radians) of the x and y axes of coordinates produced by the strain. It is to be noticed particularly that the e's are all pure numbers.

The three principal strains. It can be seen that a given uniform deformation can be described in a large number of different ways according to our particular choice of axes of coordinates. One usually works with ordinary cartesian coordinates, the three axes being at right angles, but we are still free to choose the *directions* of these axes in space, and the specification of a particular strain will vary according to this choice. For example, a shear e_{xy}, which (Fig. 26) results in an originally square element $ABCD$ deforming into the rhombus $A'B'C'D'$ can be represented equally well as a lengthening of the diagonal AC and a corresponding shortening of the diagonal BD, that is to say that a small shear e_{xy} could equally well be described as the end result of a dilatation $e_{xy}/\sqrt{2}$ along the direction OX' and a dilatation $-e_{xy}/\sqrt{2}$ along the direction OY'. One can show that *any* uniform deformation can, by a proper choice of coordinate axes, be reduced to three dilatations which are called the *principal strains*, the three shears all being eliminated, just as in the simple example mentioned. An ellipsoid with its three axes parallel to and proportional to the three principal strains is called the *strain ellipsoid*.

Non-uniform strain. We see that a uniform strain can be described in any given set of coordinates by at most *six* quantities, which are constant over the whole body. If the deformation is not uniform, we can divide the body into very small elements, each of which may be regarded as being uniformly strained, and then we can assign a set of values of the e's to each small element, which comes to the same thing as saying that the general deformation of a solid can be completely specified by a set of six strains which may vary from point to point.

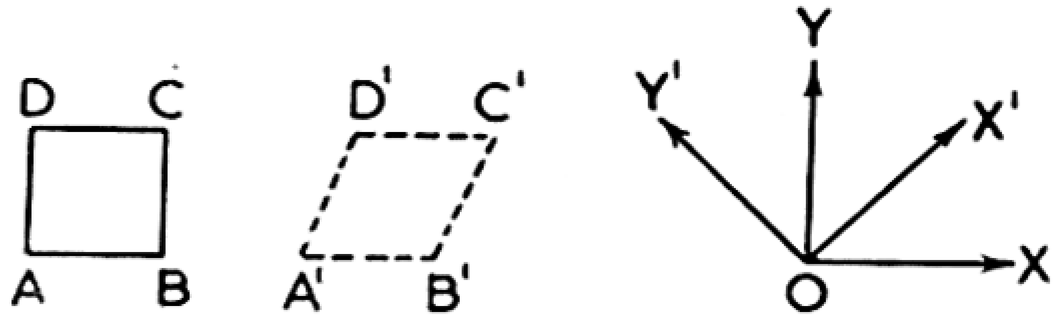


FIG. 26.

(In general, the choice of coordinate axes only enables us to remove the shears at *one* point.) Another reason for retaining the idea of a shear, even in a uniformly deformed body, is that it is very often more convenient to think of a single shear than of two separate dilatations. For example, the torsion of a wire can easily be pictured in terms of shears, but far less easily in terms of dilatations.

2. Stresses in a solid, and their effects

We now need some means of specifying the distribution of *forces* within a deformed solid at rest. We consider the forces exerted on a small element of the solid by other elements of the solid nearest to it. The forces between two small neighbouring cubes can be of two types—a mutual force along the line of centres tending to pull the cubes apart or press them together and a mutual force perpendicular to their line of centres tending to slide the cubes over one another. These two types of stress are known respectively as *tensions* or *compressions* and *shearing stresses*. In Fig. 27 some of the forces that can act on a small cubical element of the body are illustrated. If the body is at rest, the element must be in equilibrium, which means that any external forces, such as gravity, must be balanced by appropriate variations of the stresses from point to point. As a simple example of this, one may think of a heavy wire suspended vertically in which the tension increases as one goes up the wire. A second consequence of each element being at rest is that there must be no resultant couple on it, which means that forces such as X_y and Y_x, in Fig. 27 must be equal to one another, so that there are three tensions and three (not six) independent shearing stresses in the most general case. If a body is uniformly stressed the three shearing stresses can be eliminated by a suitable choice of the coordinate axes, and we can define a *stress ellipsoid* just as we defined a strain ellipsoid on p. 74. Stresses are always specified as forces per unit area, so that the force exerted in a given direction on a small cube (Fig. 27) by its neighbouring elements is equal to the corresponding stress multiplied by the area of one of the cube faces.

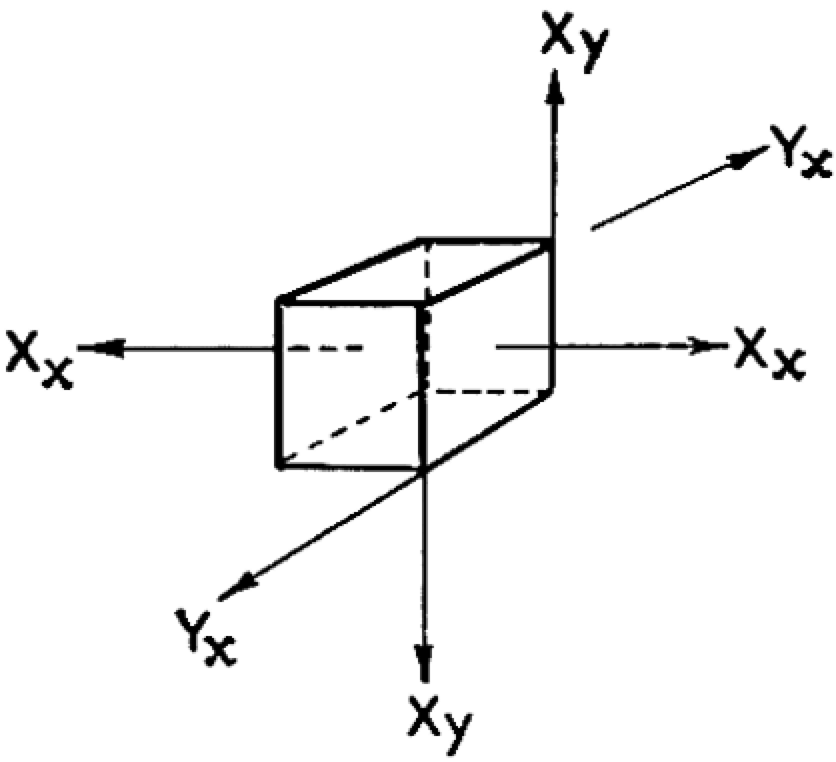


FIG. 27.

Only now that we have a means of describing numerically both the distribution of forces within a solid and the deformations that these forces produce are we in a position to compare the mechanical properties of different solids numerically. The methods of describing stresses and strains that we have just explained are of quite general application to any solid, whether it is crystalline or not. Since we have six possible strains, each of which may depend on any or all of the six possible stresses, it is clear that the number of possible types of behaviour of a solid must be very large indeed. Fortunately, there are a number of simplifying factors which enable us to classify solids into a number of broad types.

Hooke's Law.—If we imagine a small element of a solid displaced slightly from its equilibrium position, the surrounding elements being held fixed, the work done is clearly some function of the displacement of this element. We may suppose it displaced a distance x from its equilibrium position. For a small displacement, we may express the work done as a series of ascending powers of the displacement x, thus:

$$W = A_1x + A_2x^2 + A_3x^3 + \dots$$

there being no term independent of x, since $x = 0$ represents zero work done. The fact that $x = 0$ represents equilibrium also implies that A_1 must vanish and that A_2 must be positive, otherwise the potential energy would not be a minimum and the configuration represented by $x = 0$ would not be one of stable equilibrium. Thus we are led to expect, independently of the mechanism of interation between the atoms which causes the cohesive forces, that, for small displacements, the energy is proportional to the square of the displacement, that is, that the *restoring force is proportional to the displacement.* This law was enunciated by **Hooke** in the concise phrase *Ut tensio, sic vis.* Naturally, the general argument gives us no indication of the expected *range of validity* of this law, which depends on the sign and numerical magnitude of the ratio A_3/A_2. In Fig. 28 we show the relation between stress and strain in a simple case, that of a wire or bar subjected to an increasing tension. In some substances, such as steel, a number of clearly defined regions of the stress-strain curve can be discerned, in others, such as rubber, lead, or annealed copper, the division into regions is blurred or absent, while in still others, such as glass, fracture occurs at a very early stage. The curve *OABCD* represents the relation between extension and tension for a steel specimen. We shall first describe this curve, and then point out in what ways the corresponding curves for other substances behave. The region *OA* is straight, and, within it, Hooke's law is valid. A characteristic property of this region is that, if the stress is removed, the extension drops again to zero. Beyond *A*, the specimen is permanently deformed by the tension, and the point *A* is therefore called the *elastic limit.* As the tension is increased the extension increases more and more rapidly, until, at the point *B*, known as the *upper yield-point,* a finite extension rapidly occurs without any further increase of tension, and indeed, the tension needed to extend the specimen further may drop to a somewhat smaller value known as the *lower yield-point,* point *C* on the curve. The difference between *B* and *C* can be obtained only if the specimen is machined very accurately to shape, surface defects eliminated, and precautions taken both to apply the tension strictly along the axis of the specimen and to make sure that throughout the test the applied *stress* is only just sufficient to cause further elongation at any stage. If, for example, we merely hang a weight on a wire, any local "thinning

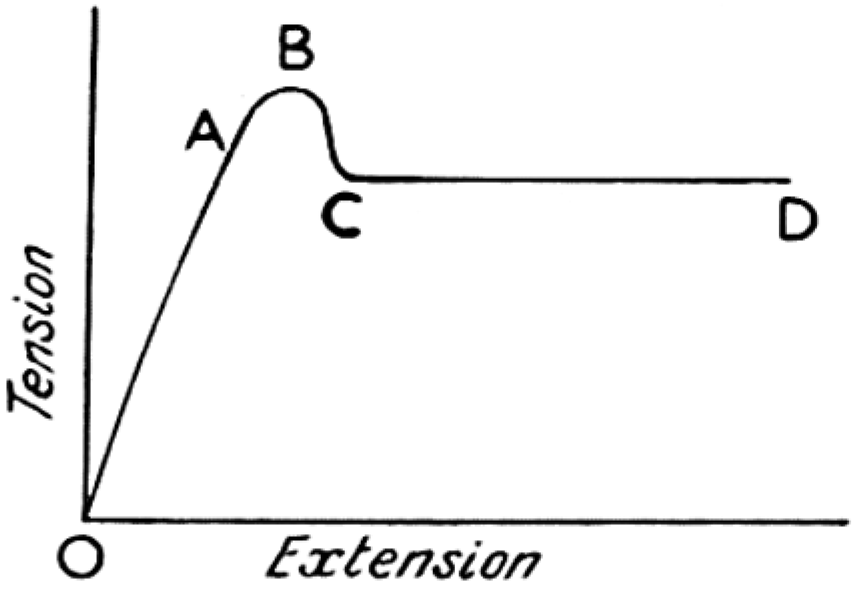


FIG. 28.

down" of the wire results in an increase in the applied *stress*, which usually results in the scale-pan running down very rapidly in the interesting region. (Engineers often describe testing machines which, *e.g.*, are designed to extend the specimen at a constant rate, the force applied being varied accordingly, as "soft", while cruder methods in which the applied *force* is constant are described as "hard".) Under ordinary laboratory conditions the hump between A and C is absent and the point at which the curve becomes flat is described simply as the *yield-point*. Beyond C, the tension per unit area required to produce further extension is practically constant, the maximum tension per unit area in this region often being known as the "ultimate" or breaking stress, and extension continues until fracture, which is usually preceded by local thinning or "necking" of the specimen. The region CD is usually known as the *plastic* region, because the extension produced is nearly proportional to the length of time for which the tension is applied, in contrast to the behaviour below the yield-point, where the final extension is attained almost immediately. The region CD usually slopes upward a little, this effect being known as *work-hardening*.

Broadly speaking, the history of any solid subjected to tension is similar to that described above. Many metals have definite elastic limits and yield-points, like steel, while for others, such as copper and lead, these quantities are very ill-defined and it is hard to distinguish between the elastic and plastic regions. Brittle substances, such as glass, are characterised by the fact that there is almost no plastic region, fracture occurring very soon after the elastic limit is passed. Rubber can be stretched to two or three times its unstretched length without permanent deformation and to eight or ten times its unstretched length without fracture, while the elastic limit for metals corresponds to an extension of less than 1 per cent. of the unstretched length, though an extension of the order of 10 per cent. of the unstretched length without fracture is often possible.

For metals, the region within which Hooke's law is valid appears to be very small (Fig. 28) but it is of extreme technical importance, because, for many engineering purposes, it would obviously not be safe to subject a metal member to stresses within the plastic range, as this would lead to permanent deformation, becoming worse each time the maximum load was applied. Indirectly, however, the engineer takes advantage of the plastic flow of metals in many ways. For example, rivets are usually inserted hot, and their contraction on cooling is used to tighten the joints. This process would be unworkable in practice but for the existence of the region of plastic flow which permits the rivets to stretch. Again, if there is a *hole* in a member subject to tensile or bending stresses, there are often regions of very high stress in the neighbourhood of the hole. This can often be ignored by the designer, for, when the load is applied, plastic flow occurs in these regions of high stress, which usually results in the stress becoming more equally distributed.

Apart from these special cases, engineers usually design structures so that the maximum stress in a member is less than the elastic limit by a certain factor, usually of the order of two or three, known as the *factor of safety*. For a member exerting *thrust*, the elastic limit has to be replaced by the "buckling load" see p. 103. For economic reasons, one tries to arrange that the factor of safety is nearly the same throughout a structure, so that the calculation of stresses in a structure is of great importance. We shall consider a few simple cases below (Art. 5).

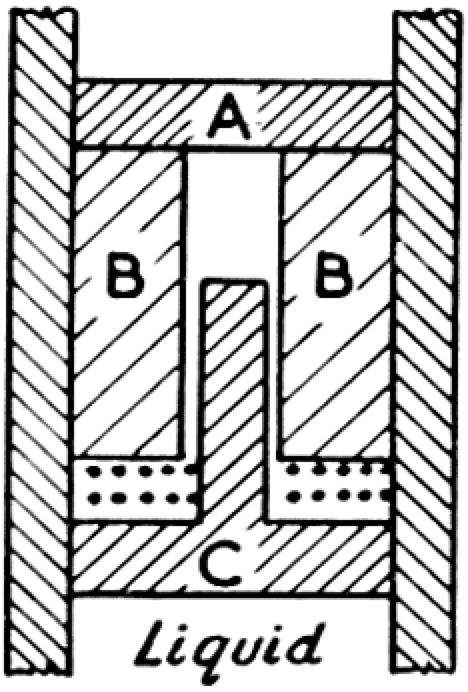


FIG. 29.

Bodies in compressive stress.—The behaviour of various bodies under pure hydrostatic pressure has been investigated by **P. W. Bridgman.** (See his book, *The Physics of High Pressures.*) By special methods he was able to reach pressures of the order of 100,000 atmospheres. One of his most important techniques is the use of a special method of packing which is leak-proof (Fig. 29). The piston *A* is worked by a press, the pressure being transmitted by the ring *B* to the piston *C*, through soft rubber packing retained in place by a copper ring. Since the total force exerted by the liquid in the pressure-vessel on *C* must be the same as that exerted by the packing, the pressure in the packing is automatically greater than in the liquid which thus cannot leak, so that attainable pressures are limited only by the strength of the containing vessel itself. Bridgman found that solids in compression first of all obey Hooke's law and usually, but not always, become less compressible as the pressure increases. The change of volume under purely hydrostatic pressure seems always to be completely reversible, there being nothing corresponding to the yield-point. In engineering practice, solids are never subjected to *pure* hydrostatic pressure, and a member subjected to a compressive force along its length, fails not by anything corresponding to plastic flow, but by *buckling*, the straight position becoming unstable (p. 103).

3. Two general theorems on elasticity

These two theorems apply not only to an elastic medium under stresses, but to any kind of framework or structure, subject only to the two conditions, first that all the members obey Hooke's law, secondly that nothing is ever stressed beyond its elastic limit (so that any work done in deforming the structure could be recovered again when it regains its unstressed configuration). We let P_1, P_2, P_3, etc. (Fig. 30), represent forces acting at various specified points in the structure, and δ_1, δ_2, δ_3, etc., the displacements of these points *along* the directions of these forces, displacements in directions perpendicular to the forces being ignored. We may, if we wish, specify separately more than one force acting at a single point, but we must then arrange matters so that the specified forces acting at a given point are at right angles to one another, as can always be done by suitable composition or resolution of the forces. Thus, in Fig. 30, δ_4 would be the displacement of point C along the line of P_4, while δ_5 would be the displacement of point C along the line of P_5, while δ_6 would be the displacement of point D along the line of P_6.

If Hooke's law holds, this means that any of the displacements can depend only linearly on the applied forces. We may fix our attention on one of the points (A, say), and imagine the forces applied to the structure one by one. The final value of δ_1 is then, by Hooke's law, a linear function of each of the forces, that is

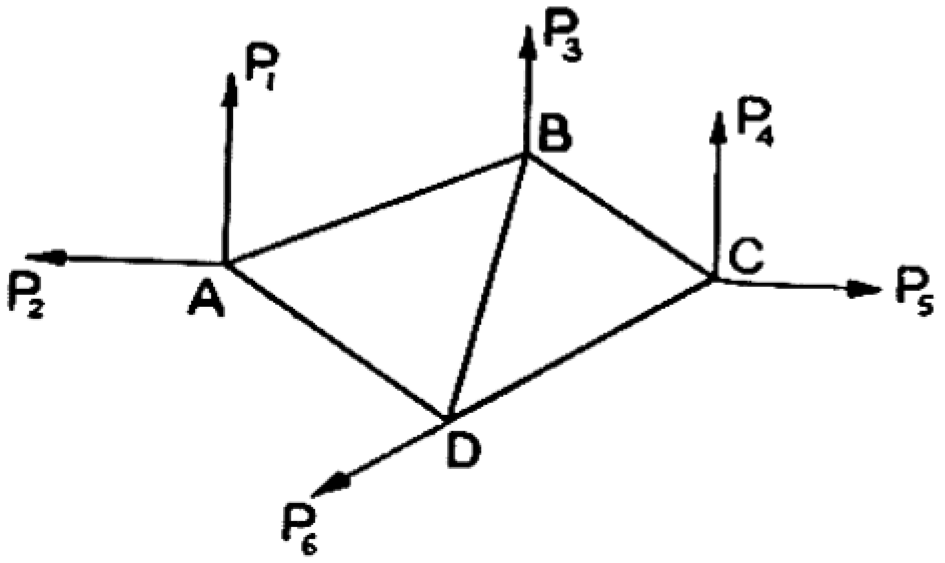


FIG. 30.

$$\delta_1 = a_{11}P_1 + a_{12}P_2 + a_{13}P_3 + \dots$$

and similarly $$\delta_2 = a_{21}P_1 + a_{22}P_2 + a_{23}P_3 + \dots$$

the quantities a_{11}, a_{12}, etc., all being constants. The first result that we shall prove is that $a_{rs} = a_{sr}$. To prove, for example, that a_{12} is equal to a_{21}, we imagine the framework unstressed, and then apply first force P_1, then force P_2. By what we have assumed, the final configuration of the framework would be the same, and the same amount of energy would be expended, if we applied first force P_2, then force P_1. If we apply P_1 first the displacement along P_1 is $a_{11}P_1$, and the work done is $\frac{1}{2}a_{11}P_1^2$, the factor $\frac{1}{2}$ arising because the displacement is, at each instant, proportional to the applied force. If we now apply P_2, keeping P_1 constant, the work done as a result of a displacement along P_2 is $\frac{1}{2}a_{22}P_2{}^2$. However, the application of P_2 may cause a further displacement along P_1, which will have the value $a_{12}P_2$, and the work done by P_1, *as a result of the application of* P_2, is then, since P_1 retains a constant value during this process, $a_{12}P_1P_2$. The total work done is then $\frac{1}{2}a_{11}P_1{}^2 + \frac{1}{2}a_{22}P_2{}^2 + a_{12}P_1P_2$. If we applied P_2 first, then P_1, the work done would be $\frac{1}{2}a_{11}P_1{}^2 + \frac{1}{2}a_{22}P_2{}^2 + a_{21}P_1P_2$, from which it follows that $a_{12} = a_{21}$. A similar argument can be applied to every other pair of coefficients. The work done as a result of applying all the forces may be written

$$\tfrac{1}{2}\sum_r\sum_s a_{rs}P_rP_s.$$

The second theorem we shall prove is very similar to Green's reciprocation theorem in electrostatics. Suppose that we have stressed the framework in two distinct ways, and that the application of the set of forces P_1, P_2, etc., has produced the displacements δ_1, δ_2, etc., while the application of another set $P_1{}'$, $P_2{}'$ . . . acting at the same points and in the same directions produces the displacements $\delta_1{}'$, $\delta_2{}'$, etc. We then have

$$\delta_r = \sum_s a_{rs}P_s, \quad \delta_r{}' = \sum_s a_{rs}P_s{}' \quad \ldots\ldots\ldots\ldots\ldots \quad (1)$$

Consider now the mixed product $\sum_r P_r\delta_r{}'$. By equation (1) this may be written $\sum_r\sum_s a_{rs}P_rP_s{}'$, or since $a_{rs} = a_{sr}$, as $\sum_r\sum_s a_{sr}P_rP_s{}'$, which, by equation (1) may be written as $\sum_s P_s{}'\delta_s$, which is the same as $\sum_r P_r{}'\delta_r$. A particular case of this result that is often useful is obtained if we suppose that all the forces vanish, except P_k and $P_l{}'$ which are equal, and we shall then have $\delta_k = \delta_l{}'$. This result is often very valuable in stress experiments, *e.g.* on models of bridges, etc., where the deflections to be measured are often small, because it enables us to reduce the number of points at which we need actually take measurements, *the downward deflection of point B by a certain weight applied at A being equal to the downward deflection of A due to applying the same weight at B.*

We may apply the first of our theorems to the problem of finding the most general possible elastic behaviour of a medium obeying Hooke's law, and not stressed beyond its elastic limit. For the forces P_1, P_2, etc., we substitute the six forces proportional to the stresses X_x, X_y, or Y_x, Y_y, etc., acting on a small cube of elastic material, the force acting on the cube being the corresponding stress multiplied by the area of a cube-face. The corresponding displacements will be the strains e_{xx}, e_{xy}, etc., each multiplied by the side of the small cube. (We recall that a strain is a dimensionless quantity measured either as an angular displacement or as a fractional extension or contraction.)

Since there are six strains, each of which may vary linearly with the six stresses, there might be as many as thirty-six coefficients of the type a_{rs} specifying the variations of stresses with strain. Our theorem that $a_{rs} = a_{sr}$ shows that this number can be reduced to at most twenty-one independent coefficients. (Six diagonal elements plus fifteen non-diagonal elements.) The actual number of independent coefficients varies according to the crystalline symmetry. For a purely isotropic substance, there are only *two* independent coefficients of elasticity.

4. The deformation of isotropic solids

It is well known that, if a wire or bar is extended, the cross-section does not remain the same, but decreases. Hooke's law leads us to expect that the fractional increase of length under tension and the fractional contraction in any direction perpendicular to the length should both be proportional to the applied tension. We can thus define two independent coefficients of elasticity of an isotropic material.

Young's modulus represents the ratio of applied tension per unit area to the fractional increase of length, supposing that the sides of the specimen are free to contract. It is usually represented by E. **Poisson's ratio** represents the ratio of the fractional lateral contraction to the fractional linear extension under tension. It is represented by σ.

There are thus at least two elastic constants needed to describe the behaviour of an isotropic body. For convenience, we often define also:

The rigidity (or torsion) modulus, usually written n, which is defined as the ratio of the applied shearing stress to the shear it produces.

The compression (or bulk) modulus which is defined as the ratio of a uniform hydrostatic pressure applied to the solid to the fractional decrease in volume that it produces.

These four quantities are not all independent, and we shall now derive two relations between them. Naturally, these four elastic constants are not the only ones that one could define; we can, for example, define a quantity analogous to Young's modulus, but with the cross-section not permitted to change its area, the sides of the specimen being rigidly held. This modulus can also be expressed in terms of any two of the others. We first express the rigidity modulus n, in terms of E and σ. To do this we consider a small cubic element of side l sheared through an angle θ, by a tangential stress, equal, by the definition of the shear modulus, to stresses $n\theta$ per unit area acting in the directions and senses $\overrightarrow{OC}$ and $\overrightarrow{BA}$ (Fig. 31). By the definition of shear stress, these correspond to forces $n\theta l^2$ on the small cube, which may be supposed to act in these directions at the mid-points of the sides of the cube through C and A perpendicular to the paper. However, these forces by themselves constitute a net *couple* on the small cube, so that such a shear stress cannot exist by itself, otherwise the small cube would not be in equilibrium.

This is a particular example of what we proved above (p. 75), namely, that there are only three independent shearing stresses (not six as we at first expect) and that $X_y = Y_x$, etc., so that the stress $n\theta$ in the X direction must be accompanied by equal stresses $n\theta$ in the Y direction. Geometrically, too, it is evident that the final shape of the cube is just the same whether, in Fig. 31, we consider the square $CDEF$ to have slid relative to $ABGH$ or whether we think of the shear as a relative sliding of the squares $BCFG$ and $ADEH$. (H is the vertex of the cube not visible in Fig. 31.) We now show that the same shear can be produced by applying a tension T per unit area along the diagonal AC and a compressive stress T per unit area along the diagonal BD. If we applied the tension along AC by itself the result would be a fractional elongation T/E along AC, and a fractional contraction $\sigma T/E$ in the two perpendicular directions, that is, both along BD and in the direction perpendicular to the paper. If we now apply an equal compressive stress along BD, the result is a further contraction T/E in this direction, together with an expansion $\sigma T/E$ along AC and an equal expansion perpendicular to the paper. Thus, the effect of applying both stresses is to leave the dimension perpendicular to the paper unaltered, and to lengthen the diagonal AC and shorten the diagonal BD by a fraction $\frac{T}{E}(1 + \sigma)$ of their

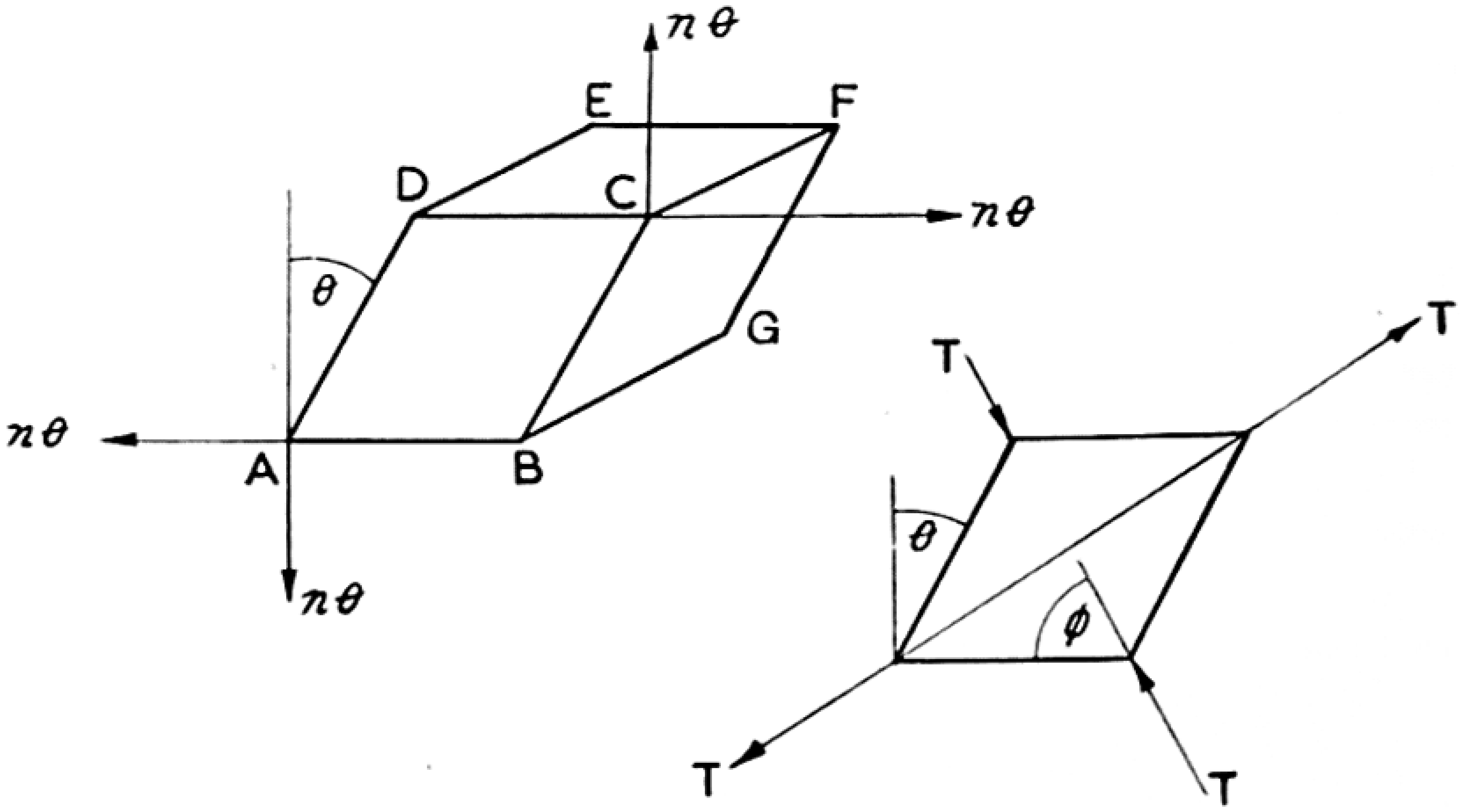


FIG. 31.

original lengths, so that we have indeed produced a simple shear in the plane of the paper without altering the dimension perpendicular to it. Fig. 31 shows that, for θ small, as is always the case for deformation in the elastic range, the length of the diagonal AC has changed from $l\sqrt{2}$ to $2l \sin \phi$, where ϕ is equal to $\pi/4 + \theta/2$ and l is the side of the unit cube. The fractional increase in the length of AC is therefore

$$\frac{2l\left[\sin\left(\frac{\pi}{4}+\frac{\theta}{2}\right)\right]-l\sqrt{2}}{l\sqrt{2}}$$

that is $\theta/2$, since for θ small we may replace $\sin \theta/2$ by $\theta/2$ and $\cos \theta/2$ by unity. Thus we have obtained the relation

$$\frac{T}{E}(1+\sigma)=\frac{\theta}{2} \quad \ldots\ldots\ldots(2)$$

by comparing the shears produced by these two types of stress, and we can obtain a relation between the elastic constants if we can obtain also a relation between the shearing stress $n\theta$ and the tensile and compressive stresses T in Fig. 31. This can be done by considering that, since these two sets of stresses produce precisely the same deformation, they must by Hooke's law, be in all respects equivalent.

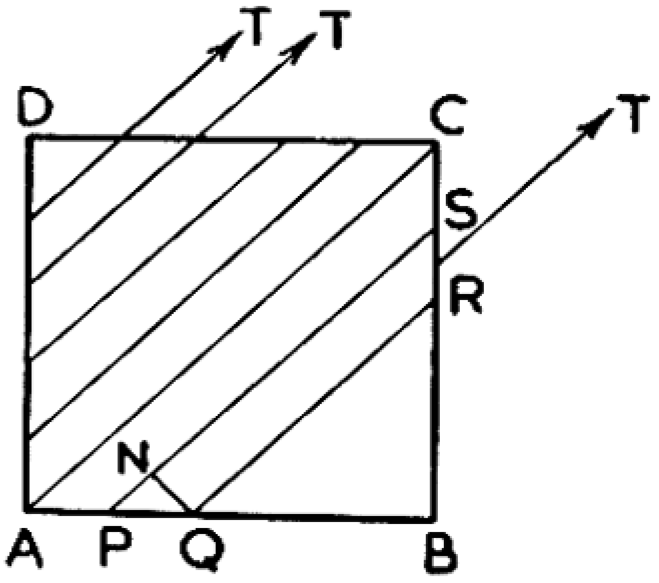


Fig. 32.

We therefore convert these stresses into forces acting on the small cube of side l, remembering that the stresses are not themselves forces, but are forces per unit area. By definition of shearing stress, a stress $n\theta$ (Fig. 31) corresponds to a force $n\theta l^2$ on the cube, which will act at the mid-point of the side of the cube CF, *perpendicular to the paper* and an equal force $n\theta l^2$ acting at the mid-point of the side of the cube AH. The total force acting at the mid-point of CF is then $n\theta l^2\sqrt{2}$, acting parallel to the diagonal AC. The total force, corresponding to the tension T acting on the cube along AC, is $T \times$ area of rectangle $BDEG$, that is, the area of the section of the cube perpendicular to T. To see this, we may imagine (Fig. 32) the triangles ABC, ADC divided into strips by lines parallel to AC, and the cube cut into slices by planes perpendicular to the plane of the paper. Consider the slice corresponding to the trapezium $PQRS$. Its area of cross-section is $QN.l$, and, since the fractional extension of all lines parallel to AC is the same, the tension acting along all these slices must be T, and the force associated with the portion RS of CB is $T.QN.l$. These forces can all be compounded into a single force $T.BD.l = Tl^2\sqrt{2}$ acting parallel to AC through the mid-point of the side of the cube through C perpendicular to the paper. Similarly, we can show that a force $Tl^2\sqrt{2}$ acts through the mid-point of the side through A perpendicular to the paper, and the compressive stresses parallel to the diagonal BD are equivalent to equal forces acting through the mid-points of the sides through B and D perpendicular to the paper. Comparing the two sets of forces required to produce the same shear, we have, finally, $T = n\theta$. Substituting this result in equation (2) we have finally

$$E = 2n\,(1 + \sigma) \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (3)$$

This result has taken some time to establish because we did not know the law for resolution of stresses corresponding to that for the resolution of forces. Stresses, although they have magnitude and direction, are *not* vectors, but are *second-order tensor* quantities, and transform according to a different law. For fuller discussions of this point, and for a proof of the statement on p. 75, that any uniform stress can be regarded as consisting of not more than three tensions or compressions at right angles to one another, the reader is referred to advanced treatises on elasticity such as those by Love or Southwell.

We now establish a relation between Young's modulus and the bulk modulus. This can be done quite simply as follows. We imagine an isotropic body subjected to three equal compressive stresses P per unit area, parallel to the three coordinate axes. If we again think of a small cube of the material of side l, the forces Pl^2 on each cube face are just those that would arise if a uniform pressure P were applied to the body. Suppose that, as a result, the lengths of the sides change by dl, the volume changes by $(l + dl)^3 - l^3$ or $3l^2dl$, if dl is small. Thus, by the definition of the bulk modulus K, we have

$$P = 3Kl^2dl \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (4)$$

We now imagine these three stresses applied one at a time. The effect of applying a compressive stress P along the X axis is to shorten the cube by lP/E along this axis, and to lengthen it by $\sigma lP/E$ along the Y and Z directions. Similarly, when we apply equal stresses along the Y and Z axes, the cube is, in each case shortened along the direction of the applied stress, and lengthened in the other two directions, the final change in length along any axis being $lP/E - 2\sigma lP/E$ which we may now put equal to dl in equation (4). We thus have

$$E = 3K\,(1 - 2\sigma) \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (5)$$

Equations (3) and (5) represent two equations between the four elastic constants E, K, n, and σ that we have defined for an isotropic body, so that, theoretically, only two of them are actually independent, and a measurement of two of them should be sufficient to determine the other two. Equations (3) and (5) also show us that Poisson's ratio σ must lie between -1 and $+\frac{1}{2}$, otherwise either the rigidity modulus or the bulk modulus would become infinite. In practice, no substance with a negative Poisson's ratio is known, and, indeed, one would hardly expect a body to expand laterally if it is being extended, which is what a negative Poisson's ratio would mean.

Thus, in practice, the possible values of σ range between 0 and $\frac{1}{2}$. A value of $\frac{1}{2}$ would imply that the material was incompressible. For most substances σ is in the region of $\frac{1}{3}$. Students are often puzzled by the fact that, with *wire* specimens in the laboratory, independent measurements of E, σ, and n often give values quite inconsistent with equations (3) and (5), although the measurements themselves seem to be quite satisfactory and reproducible. This happens because the process of wire-drawing usually leads to some alignment of the minute crystals in the metal, so that they are no longer oriented completely at random This may result in the metal becoming non-isotropic, in which case equations (3) and (5) need not apply, and the wire really needs more than two independent elastic constants to describe its behaviour properly.

5. Simple examples of stress-calculations

Having established the laws of deformation of elastic isotropic solids, one is then led to try to take a further step, namely, to calculate the deformation of a given elastic body when one is given, not the stresses at each point, but simply its elastic constants and the details of the applied load—that is, we have to find a distribution of stress and strain in the body that is consistent both with the laws of elasticity and the shape and size of the body, that is, we have to solve the equations of elasticity subject to certain *boundary conditions*.

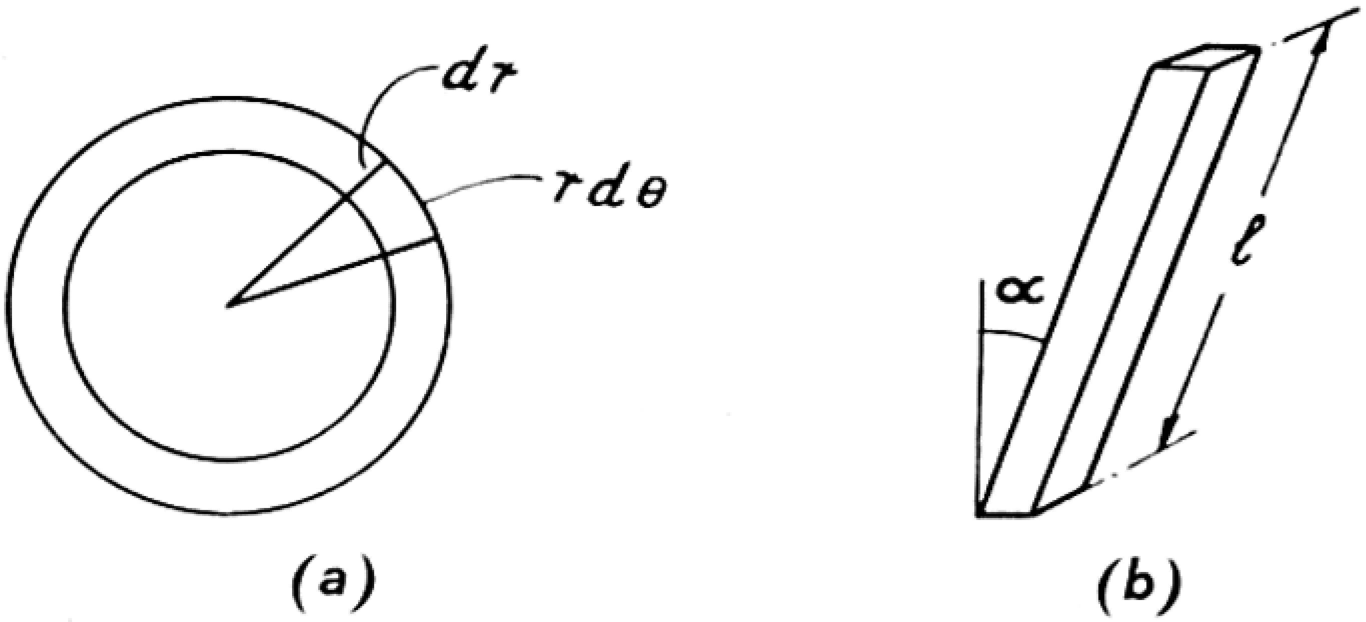
This type of problem is a familiar one in electrostatics, where the problem is to find a distribution of potential V that is constant on certain conductors and satisfies the equation of Laplace

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} = 0.$$

(See Jeans, *Electricity and Magnetism*.) Most problems in elastic deformation can be reduced to the solution of an equation

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right)^2 V = 0.$$

Functions that satisfy this equation, but not Laplace's, are usually called **biharmonic**, while solutions of Laplace's equation are called harmonic functions, so that the solution of an elastic problem involves some combination of harmonic and biharmonic functions. Exact solutions of many problems can be obtained in this way, just as in electrostatics and hydrodynamics, provided that the geometry is fairly simple, and such solutions are very valuable in some problems, for example in determining the increase of stress resulting from cutting a hole in a girder or deck. However in determining the stresses to be expected in an actual structure under given loading one cannot hope for an *exact* solution for the whole structure and each member has to be considered separately, the stresses and strains in members that are joined together being adjusted (by a process of successive approximation) to be as consistent with one another as possible, this process being known colloquially as "stressing" the structure. We shall discuss in this book the deformations of a few simple types of member, the results for which are needed in the experimental determination of elastic constants.



(a) (b)

FIG. 33.

(*a*) *Torsion of a cylindrical rod.*—We may imagine a circular rod of length l and radius a, twisted through a total angle ϕ. We suppose the rod divided up into cylindrical shells and consider a slice of one of these shells of inner and outer radii r and $r + dr$ lying between the polar angles θ and $\theta + d\theta$ [Figs. 33(*a*) and (*b*)]. Fig. 33(*a*) represents the base of such a slice of area $rdrd\theta$, while Fig. 33(*b*) represents its side elevation. For a uniform twist, the originally straight sides remain straight, and the slice in question has been deformed by a simple shear. Since the two ends of the rod have rotated through a relative angle ϕ, the top and base of the slice have been displaced by a distance $r\phi$, that is, the slice has been sheared through an angle $\alpha = r\phi/l$. By the definition of shear modulus, a tangential force $n\alpha \times$ area of base of slice is required to produce the deformation α, that is a force $n \,.\, r\phi/l \,.\, rd\theta dr$. The moment of this force about the axis of the cylinder is obtained simply by multiplying this by r. Integrating over the whole cross-section of the rod, we obtain for the couple that must be applied at each end to maintain the twist the value:—

$$G = \int_0^a \int_0^{2\pi} \frac{nr^3\phi}{l}\, d\theta\, dr = \frac{n\pi a^4 \phi}{2l} \quad \ldots\ldots\ldots\ldots \quad (6)$$

For a hollow cylinder of internal radius a_i it is evident that a^4 in equation (6) must be replaced by $a^4 - a_i^4$.

The treatment of torsion of a body that has any section other than circular is very much more difficult, because plane sections do not in general remain plane. Only comparatively simple shapes of cross-section can be treated at all, and the reader is referred to Love, *Theory of Elasticity*, for details. One result is worth quoting. For an elliptic cross-section of semi-axes a and b, **St Venant** showed that $a^4/2$ in equation (6) must be replaced by $a^3b^3/(a^2+b^2)$, and, if b is very small compared with a, so that the body is a strip with rounded corners, this becomes effectively equal to ab^3. Now, if we imagine a circular wire of radius a flattened into such a strip without change of cross-section, the semi-axes being a_1 and b_1, we should have $a_1b_1 = a^2$. The couple produced by torsion through a given angle is changed in the ratio $a_1b_1^3/a^4 = b_1^2/a^2$. This fact is often made use of in designing galvanometer and similar suspensions where one wants the minimum restoring couple consistent with the suspension having enough tensile strength (the latter quantity being proportional to the cross-sectional area).

(*b*) *Bending of a beam.*—We imagine a beam, not necessarily of uniform cross-section, subjected to some form of transverse loading including possibly the effect of the weight of the beam itself. We usually distinguish between two different kinds of support, first "simple support", in which the beam merely rests on the support, which thus exerts a force on it, secondly, "clamping" or "building-in", in which the beam is held securely for a finite part of its length, *e.g.* by a clamp, or by passing through a hole in a wall. A "simple" support can exert nothing but a force on the beam whereas a "built-in" support can exert a couple on the beam as well.

Given a certain distribution of external load, we have to determine the deformation, and the first step is to determine the forces and couples acting at the supports, which can be done in the usual way by considering the equilibrium of the whole beam. We now consider (Fig. 34), a beam on two supports P and Q, loaded in any manner, and consider the equilibrium of the portion of the beam to the right of the plane AB. The total transverse force acting between the two portions of the beam, the so-called *shearing force* (not to be confused with the shearing *stress*, which is the force per unit area of cross-section) is simply the resultant of the forces due to any loading on the right-hand portion and the force at the support Q. There may also be a couple between these two portions of the beam, the *bending moment*, which, again considering the equilibrium of the right-hand portion, must have a moment equal to the algebraic sum of that of the force at the support at Q and of the moments of any loading on the right-hand portion, the moments being taken about an axis through the mid-section of the beam and the plane AB. Consider a uniform beam of length $2a$, supported at points P and Q distant b from each end, and with a load w per unit length. If the position of AB is x, measured from the left-hand end, the forces at P and Q are both equal to wa, while the load on the portion to the right of AB is $w(2a - x)$, acting at its mid-point, that is, at a distance $a - x/2$ from AB. The shearing force at AB is thus equal to

$$\begin{array}{rl} 2wa - w(2a - x) & x < b \ (AB \text{ to the left of } P) \\ wa - w(2a - x) & b < x < 2a - b \ (AB \text{ between } P \text{ and } Q) \\ -\,w(2a - x) & x > 2a - b \ (AB \text{ to the right of } Q) \quad \ldots\ldots \quad (7) \end{array}$$

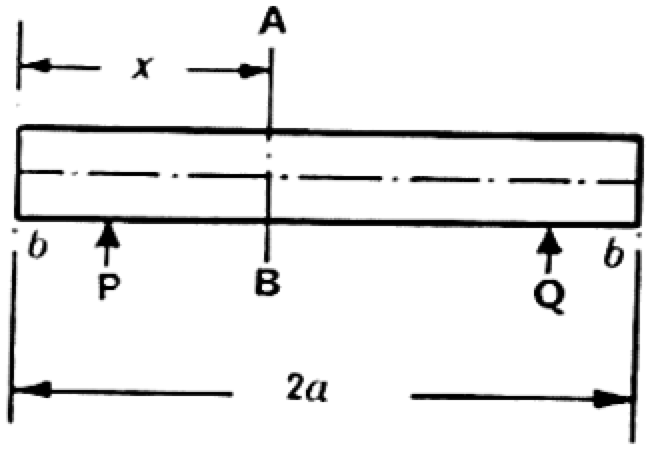


FIG. 34.

Similarly, we get the following values for the bending moment

$$\begin{array}{rl} wa(b - x) + wa(2a - b - x) - w(2a - x)(a - x/2) & x < b \\ wa(2a - b - x) - w(2a - x)(a - x/2) & b < x < 2a - b \\ -\,w(2a - x)(a - x/2) & x > 2a - b \quad \ldots\ldots\ldots \quad (8) \end{array}$$

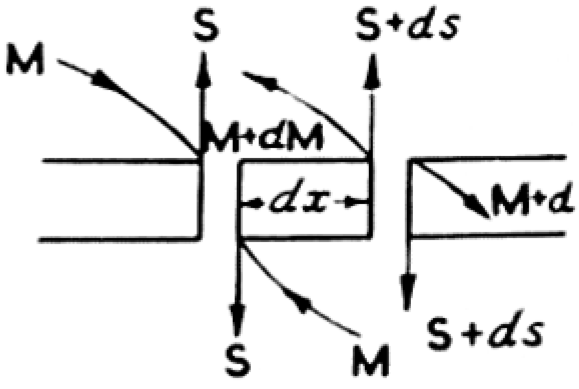


FIG. 35.

The calculations are easily extended to cover cases where concentrated loads are present, simply by adding the appropriate terms to the bending moment and shearing force, or where the loading w per unit length of beam is not a constant. In the latter case the length of the beam to the right of AB is divided up into small elements, and the total contribution to the bending moment and shearing force can be found by integrating along the length of the beam to the right of AB. In this way we can establish general relations between the load per unit length w, the shearing force S, and the bending moment M. In Fig. 35, let the shearing force and bending moment exerted on a small element of length dx by the portion of the beam to the left of it be S and M respectively, while $S + dS$ and $M + dM$ are the shearing force and bending moment exerted on the element by the portion of the beam to the right of the element. We make the sign convention that a positive shearing force acts upwards on the portion of the beam to the left of it, and that a positive bending moment tends to turn this portion of beam in a counter-clockwise sense. The net external force acting on the element dx is wdx, which must be balanced by the change of shearing force along the length of the element so that

$$dS = wdx \quad \text{or} \quad w = \frac{dS}{dx} \quad \dots\dots\dots\dots\dots\dots \quad (9)$$

The two shears S, acting at opposite ends of the element in opposite senses, are equivalent to a positive couple Sdx, which must be balanced by the difference dM between the moments exerted on the element by the portions of the beam to the right and left of it, that is:—

$$dM + Sdx = 0 \quad \text{or} \quad S = -\frac{dM}{dx} \quad \dots\dots\dots\dots \quad (10)$$

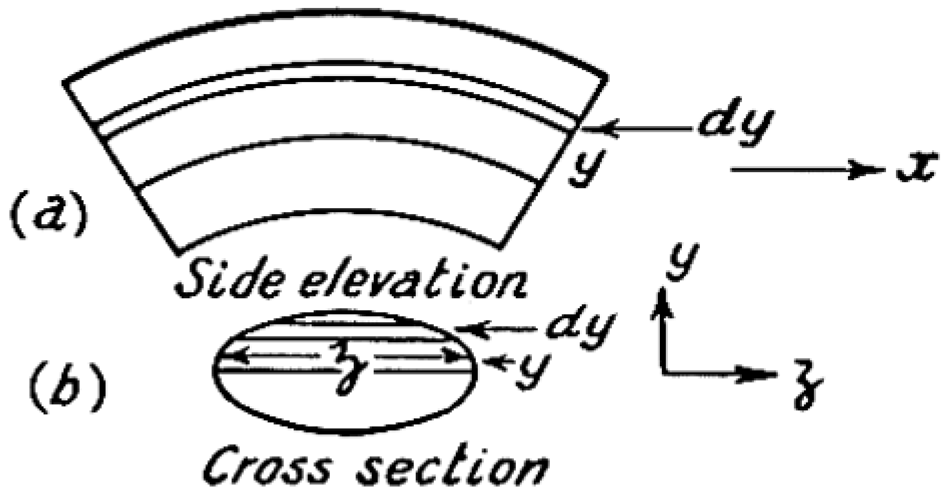


FIG. 36.

We now have to calculate the deformation due to these stresses. In nearly all practical cases, the deformation due to the shear S is small compared with that due to the bending moment M, and we shall therefore calculate the latter first. If a bar is bent there will always be a surface, the *neutral surface*, in the neighbourhood of which the fibres are neither stretched nor compressed. The part of the bar outside the neutral surface has to extend, while that inside has to be compressed in order that the bar may bend. Fig. 36(a) represents a side view of a portion of the bar. Suppose the local radius of curvature is ρ, and consider a filament parallel to the neutral surface and distant y from it, subtending an angle $d\theta$ at the centre of curvature. The length of this portion is $(\rho + y)\, d\theta$, while the length of the corresponding portion of the neutral surface is $\rho d\theta$, so that the fractional extension at a height y above the neutral surface is y/ρ, and the force due to the extension of the filament is Ey/ρ times the cross-section of the filament, which Fig. 36(b) shows to be given by $\frac{Ey}{\rho}zdy$. The contribution of this filament to the bending moment about a line in the neutral surface is therefore $\frac{Ey}{\rho}yzdy$, so that the bending moment is given by the expression $E\int zy^2dy/\rho$, the integration being taken over the whole area of cross-section, the fact that the filaments below the neutral surface are in compression and those above it in tension being automatically taken care of by taking the origin of y in the neutral surface. The integral $\int zy^2dy$ is often called the *geometrical moment of inertia of the cross-section* because it is proportional to the mechanical moment of inertia of a lamina of the same shape as the cross-section. It is often written J, and the product EJ is called the *flexural rigidity*. Now if $p\ (x)$ specifies displacement of the neutral surface from its equilibrium position at the point x on the beam, that is, the curve into which the beam is bent, we have

$$\frac{1}{\rho} = \frac{d^2p}{dx^2} \Big/ \left(1 + \left(\frac{dp}{dx}\right)^2\right)^{\frac{3}{2}},$$

but this is nearly equal to d^2p/dx^2 as long as the radius of curvature ρ is large. We then have

$$M = EJ\frac{d^2p}{dx^2} \quad \dots\dots\dots\dots\dots\dots \quad (11)$$

which gives us a relation between the bending moment and curvature of the beam. Equations (9), (10), and (11) are theoretically sufficient to enable us to calculate w if p and J are known as functions of x, while if w is known, p can be calculated if J is a sufficiently simple function of x to enable these equations to be solved. If the cross-section is uniform, J is constant, and we then have

$$w = -EJ\frac{d^4p}{dx^4} \quad \dots\dots\dots\dots\dots\dots \quad (12)$$

which can be solved by repeated integration for any given type of loading.

We have, however, neglected the deflection due to the shear. If we consider the small element dx of the beam (Fig. 35), and suppose the cross-section of the beam is A, the shearing *stress* is S/A and the element is deformed through an angle S/nA, so that the inclination of the beam dp/dx is increased by Sdx/nA as one travels from the point x to the point $x + dx$, that is to say

$$\frac{d^2p}{dx^2} = \frac{S}{nA} \quad \dots\dots\dots\dots\dots\dots \quad (13)$$

We now compare the deflections due to bending moment alone and shear alone for a representative example, namely, a beam loaded uniformly along its length, so that we may put $w = w_0$, and simply supported at its two ends, which we may take at $x = 0$ and $x = 2a$. To calculate the deflection due to the bending moment alone, we have from equations (9) and (10)

$$M = -\iint w dx dx = -\frac{w_0 x}{2}(x - 2a)$$

since the constants of integration must be chosen to make M vanish at the supported ends. (We cannot, as one might think at first, make S vanish at the supported ends, because of the forces exerted on the ends by the supports.) Integrating equation (11) twice, and using the fact that the deflection p itself must vanish at both supports, we obtain for the final profile the result

$$p = -\frac{w_0}{24EJ}(x^4 - 4ax^3 + 8a^3x) \ldots\ldots\ldots\ldots \quad (14)$$

which corresponds to a deflection of the middle of the beam ($x = a$) of $\frac{5w_0a^4}{24EJ}$, due to the bending moment alone.

We now calculate the deflection that would be produced by the shearing force S alone if the bending moment were not present. From equation (9), for uniform loading w_0, we have

$$S = w_0 x - w_0 a,$$

the constant of integration having been chosen so that S is numerically equal to the force $w_0 a$ on the supports at each end. (The difference in sign at $x = 0$ and $x = 2a$ is due to the fact that, according to the convention we have made, illustrated in Fig. 35, a shearing force is reckoned as positive at a given point when the force acts upwards on the portion of the beam to the left, and it is therefore of opposite sign at the two ends of the beam.) Inserting this value in equation (13), integrating twice, and using the fact that p must vanish at $x = 0$ and $x = 2a$, we obtain for the deflection due to shear only

$$p = \frac{w_0}{6nA}(x^3 - 3ax^2 + 2a^2x) \ldots\ldots\ldots\ldots\ldots \quad (15)$$

This vanishes at the middle of the beam, and is anti-symmetrical about the middle, just as is the shear itself. It is in any case unimportant compared with the deflection due to M. Its maximum value occurs in the neighbourhood of $x = a/2$, and is approximately equal to $\frac{w_0a^2}{16nA}$.

The ratio of this to the deflection due to M is therefore of the order of magnitude

$$\frac{a^3}{16nA}\Big/\frac{5a^4}{24EJ} \quad \text{or} \quad \frac{0{\cdot}3E}{n}\,\frac{J}{Aa}.$$

Since Poisson's ratio σ is always less than $\frac{1}{2}$, E/n can never be greater than 3 for an isotropic material. J, the moment of inertia of the cross-section, is not larger than A, the area of the cross-section multiplied by b, the maximum dimension of the cross-section. The ratio of the two types of deflection is therefore never as large as b/a, and is therefore small, because one cannot describe a body as a "beam" unless its length is large compared with its other dimensions. A similar result holds however the beam is supported or clamped, and we may therefore neglect the deflection due to shear.

It is possible to carry out similar calculations when the beam is non-uniform or is held in other ways than by simple support of its ends, equations (9) and (11) being quite general, any non-uniformity of the beam being taken care of by a variation of J with x, while the effect of different kinds of support at the ends can be taken account of by means of the constants of integration arising when these equations are integrated. These constants have to be chosen to satisfy the boundary conditions at the two ends. For convenience, we may list these as follows:—

(i) At a *simply-supported* end $p = 0$ and M or $d^2p/dx^2 = 0$, but $S \neq 0$.

(ii) At a *clamped* end, $p = 0$ and $dp/dx = 0$, but M and $S \neq 0$.

(iii) At a *free* end, M and $S = 0$, but p and $dp/dx \neq 0$.

In all cases the axis of p is taken in the direction of the unloaded beam. If w and J are given functions of x, integration of equations (9) to (11) (p. 89) introduces four arbitrary constants and, at each end, there are two boundary conditions to be satisfied.

Limitations of the simple theory of bending.—In the development of the above theory we have expressed or implied various assumptions which we shall now examine.

(1) *That the cross-section of the beam is unaltered.*—Since the portions of the beam above the neutral surface are stretched, there must be a corresponding lateral contraction σ times as great, while there must be a lateral extension below the neutral surface. The cross-section of a rectangular beam bent so as to be convex upwards along its length is therefore as shown in Fig. 37, and is concave upwards across its length, so that it has opposite curvatures in the two directions, like a saddle

or col. Comparing Fig. 37 with Fig. 36(*a*), it is evident that as a filament distant y above the neutral surface is extended a fractional amount y/ρ, where ρ is the radius of curvature along the length, but contracts laterally a fractional amount $\sigma y/\rho$, the radius of curvature ρ across the length is given by ρ/σ. There is a slight alteration in J as a result of this change in shape of cross-section due to the bending, but this can almost always be neglected.

(2) *That the radius of curvature is large compared with the thickness.*—This is almost always satisfied for elastic bending, because, as we have seen, the elastic limit for metals is exceeded if the deformation is as much as 1 per cent. anywhere.

(3) *That the minimum deflection of the beam is small compared with its length.*—This is usually satisfied in engineering practice, but exceptions occur, *e.g.* for a very thin body such as a clock-spring, which can be deflected very considerably without the elastic limit being exceeded. This case can be dealt with if we take account of the fact that $1/\rho$ is not strictly equal to $\frac{d^2p}{dx^2}$ but to $\frac{d^2p}{dx^2}\Big/\left(1+\left(\frac{dp}{dx}\right)^2\right)^{\frac{3}{2}}$ so that equation (11) (p. 90) has to be replaced by

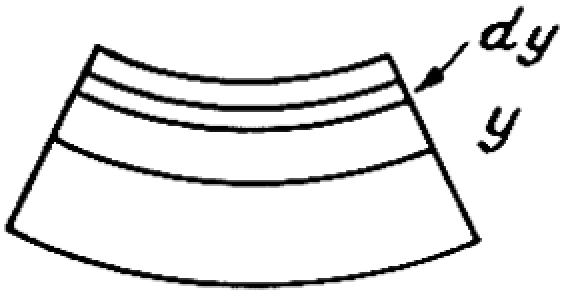


FIG. 37.

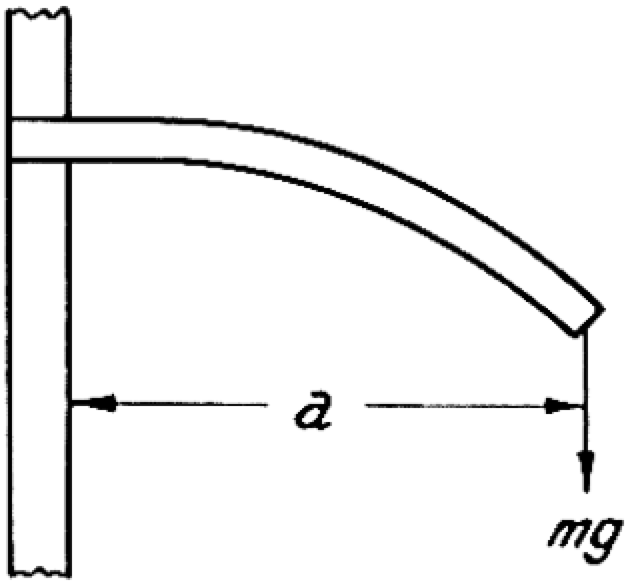


FIG. 38.

$$\frac{d^2p}{dx^2}\Big/\left(1+\left(\frac{dp}{dx}\right)^2\right)^{\frac{3}{2}} = M/EJ.$$

If M is a given function of x, this equation can be integrated once to give

$$\frac{dp}{dx}\Big/\left(1+\left(\frac{dp}{dx}\right)^2\right)^{\frac{1}{2}} = \frac{1}{EJ}\int M\,dx \qquad \ldots\ldots\ldots\ldots \quad (16)$$

As an example, consider a light cantilever "built-in" to a wall and loaded with a concentrated mass m at the free end (Fig. 38). Taking the wall as the origin, we have $M = mg\,(a - x)$, where a is the coordinate of the free end. Since $dp/dx = 0$ at the wall, equation (16) becomes

$$\frac{dp}{dx}\Big/\left(1 + \left(\frac{dp}{dx}\right)^2\right)^{\frac{1}{2}} = \frac{mg}{EJ}\left(ax - \frac{x^2}{2}\right) \quad \ldots\ldots\ldots \quad (17)$$

which can be integrated a second time in terms of elliptic integrals. Curves, satisfying equations such as (16), are often known as **elasticas,** but a determination of their analytic form is often very complicated, although certain results can be obtained by quite simple means. Suppose, for example, that we have a clock spring clamped at one point and loaded at the free end. As we pay out more and more of the spring through the clamp, keeping the load constant, the free end drops and finally becomes vertical. The horizontal distance a of the free end is then a maximum, and, if more of the spring is paid out, part of it will hang vertically. At $x = a$, we shall then have dp/dx infinitely great, so that the left-hand side of equation (17) becomes equal to unity, and we shall have $1 = mga^2/2EJ$.

(c) *The flat spiral spring.*—By a "flat" spiral spring one means a wire wound into a uniform helix of *small pitch*. Suppose that r is the radius of the wire and R the radius of the helix of N turns. If the spring is extended the effect is to impose a *twist* on the wire, for if we imagine every element of the wire to be twisted through an angle θ per unit length, the distance between the beginning and end of each turn is increased by $2\pi R \,.\, R\theta$, so that the total increase x in length is $2\pi N R^2\theta$. If a mass m is hung at the centre of the spring, it exerts a moment mgR which acts on every element of the wire, and is balanced by the torsional couple produced by the twisting, so that we have, by equation (6) (p. 87)

$$mgR = \frac{n\pi r^4\theta}{2}.$$

Hence
$$x = \frac{4mgR^3N}{nr^4} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (18)$$

Thus, *extending* the spring produces a *twist* in the wire, and the *extension* is governed by the *rigidity* modulus n.

Next, we imagine couples C applied to the top and bottom of the spring, acting round the axis of the helix. The couple required to change the curvature of the spring from $1/R$ to $1/(R - \delta R)$ is, by the same kind of argument that led to equation (11) (p. 90) for the bending of a straight beam, given by

$$EJ\left(\frac{1}{R - \delta R} - \frac{1}{R}\right) = C.$$

For the *mechanical* moment of inertia of a circle of radius r about a diameter Routh's rule (p. 31) gives $Mr^2/4$, and the *geometrical* moment of inertia J is obtained by replacing M by the area πr^2 of the cross-section. We thus have

$$C = E\frac{\pi r^4}{4} \cdot \frac{\delta R}{R^2}.$$

Now, if the radius of each coil of the spring is reduced by δR, the length of the whole spring being kept constant, the angle θ, through which the ends of the spring turn relative to one another, is

$$\theta = 2\pi N\frac{\delta R}{R},$$

where N is the number of coils. We thus have

$$C = \frac{Er^4}{8NR}\theta \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (19)$$

which means that *twisting* the spring involves *bending* the wire, which is governed by Young's modulus E.

6. Measurement of the elastic constants of solids

Almost any of the equations of Art. 5 above can be made the basis of a measurement of the appropriate elastic constant. For E and n a large number of methods are available, of which a few will be described. They are either static, a known stress being applied by means of weights, and the small deflection measured, or dynamic, the oscillations of a body of known mass or moment of inertia being timed.

(*a*) *Measurements of E by stretching.*—For a wire, one of the best methods is due to G. F. C. Searle. Two wires, one serving as a reference, are hung from the ceiling or other support, and are kept taut by means of weights. The load on one wire is never altered, but that on the other is progressively changed. The wires are connected by a light metal framework (Fig. 39) carrying a spirit-level L. The relative movement of the wires as the load on one of them is altered is compensated by a spherometer wheel S, which is adjusted after each alteration of load until the bubble is again between the marks. It is usually necessary to load and unload a few times before reproducible readings are obtained. (The idea of using a second wire as reference is to correct for any yield of the support as the load is altered.)

For a bar, the load is usually applied by a testing machine and the extension measured by a device such as *Ewing's extensometer* (Fig. 40), which is simply a means of magnifying the small extension of the bar, the elongation of the bar between A and B carrying the scale C past the microscope D.

(b) *Measurements of E by bending.*—From equation (11) (p. 90) we see that the radius of curvature of a beam is (apart from the small deflection due to shear), inversely proportional to the couple applied at any point of the rod. Now, if a beam is loaded as in Fig. 41 with two equal loads outside two supports, there is no shear anywhere between the supports, and the bending moment is constant, being equal to the sum of the two couples acting on the beam. The beam is therefore bent into an arc of a circle. J can be found if the cross-section of the beam is known, and the problem is therefore that of measuring the radius of curvature of the portion of the beam between the supports.

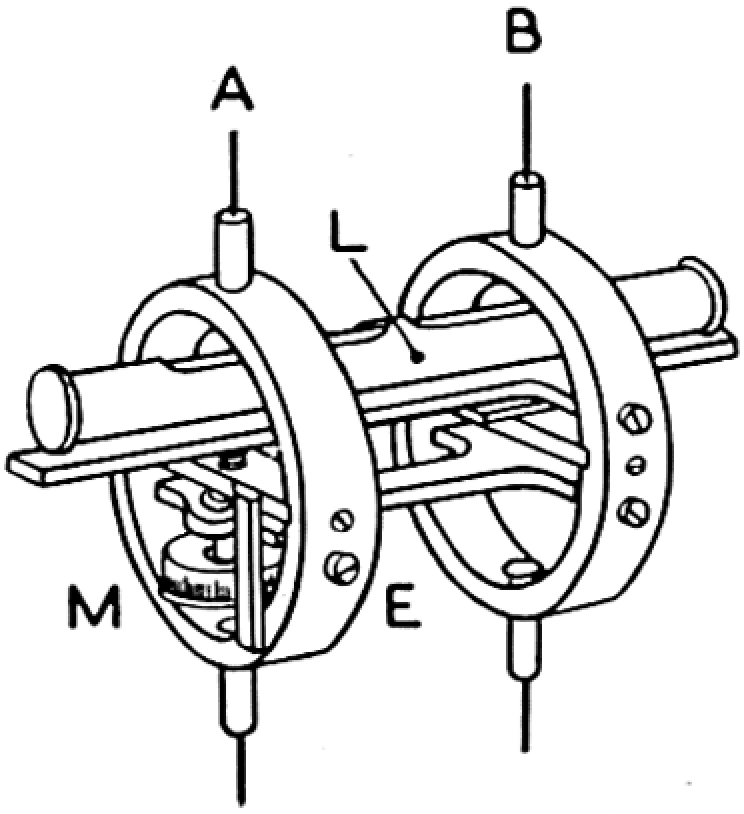


FIG. 39.

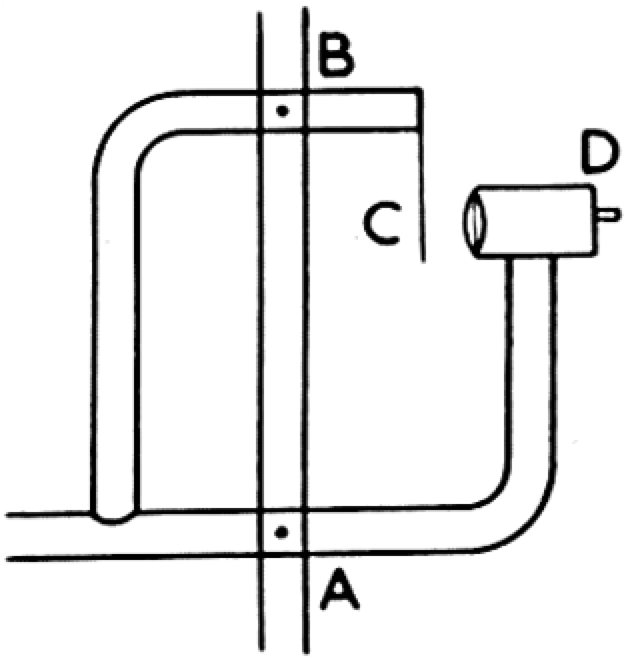


FIG. 40.

(The portions of the beam *not* between the supports are, of course, not subjected to a uniform bending moment.) This radius can either be measured directly, or else inferred from a measurement of the deflection of the centre of the beam. As a method of direct measurement, one may fix mirrors M, M' on supports a known distance apart a, and the effect of bending the beam into an arc of radius ρ is to tilt the mirrors through an angle a/ρ relative to one another, which angle can be measured easily by focussing the image of a galvanometer light on a scale, after reflection by both mirrors. Indirectly, we can determine ρ if we can measure p_0, the deflection of the centre of the beam, for if b is the distance between the two supports S we have by the geometry of a circle

$$p_0 (2\rho - p_0) = \frac{b}{2} \cdot \frac{b}{2}.$$

p_0 can be measured by any method suitable for a small distance, such as a spherometer or by interferometry, the latter method being capable of measurement to an accuracy of a fraction of a wave-length of light (see **Cornu's** method below). Another method due to **Whiddington** is by measuring, by a beat-frequency method, the small change in capacity of an electrical condenser of which the beam forms one plate, due to its motion relative to a fixed plate.

(c) *Measurement of E by dynamic methods.*—The easiest way of measuring E for a specimen in the form of a thin strip is to clamp it securely, attach a suitable weight to the free end, and time the flexural oscillations of the resulting system (Fig. 38). For if we measure x from the clamped end, the force P, that has to be applied at the free end to hold it at a small distance l below its equilibrium position, may be found as follows. If we measure from the clamped end as zero, and the length of the bar is l, the bending moment at the point x, due to P, is $P\,(l-x)$. The deflection of the beam is then given by equation (11) (p. 90) as

$$\frac{d^2p}{dx^2}=\frac{P}{EJ}\,(l-x).$$

Integrating twice, and using the fact that p and dp/dx vanish at $x=0$, we have

$$p=\frac{P}{EJ}\left(\frac{lx^2}{2}-\frac{x^3}{6}\right)$$

so that the deflection p of the end of the beam ($x=l$) requires a force

$$P=\frac{3EJp}{l^3},$$

which is the force acting on the mass m. We thus have the net force on this mass

$$\frac{3EJp}{l^3}-mg$$

so that if the end is displaced a small distance q from its equilibrium position after m has been added, the extra force acting on the mass is

$$\frac{3EJ\,(p+q)}{l^3}-\frac{3EJp}{l^3}$$

and the corresponding acceleration, neglecting the mass of the spring compared with m, is $\frac{3EJq}{ml^3}$, so that the equation of motion is

$$\ddot{q}+\frac{3EJq}{ml^3}=0,$$

and the corresponding period is given by (p. 18)

$$T = 2\pi\sqrt{\frac{3EJ}{ml^3}}.$$

(*d*) *Measurements of* σ.—By the definition of σ it is necessary to measure the lateral contraction, that occurs during a Young's modulus experiment, to determine σ as well as E. In the apparatus shown in Fig. 39 the lateral contraction of a fairly thick wire can be measured by a screw-gauge, or if the specimen is in the form of a bar, the lateral contraction can be measured by a method due to **Morrow.** The specimen is gripped laterally by a spring caliper and the relative motion of the two arms is magnified by a mirror and "optical lever". Again, using the apparatus shown in Fig. 41, it was pointed out above (p. 92) that if a beam were bent, it acquired an opposite curvature in the direction at right angles to the axis of bending. Thus the curvature along the breadth of the beam can be measured by the two-mirror principle just described (p. 96). As an example of an interferometric method, we may mention that due to **Cornu,** which is particularly valuable for a substance such as glass, to which only very small stresses can be applied with safety. The specimen is subjected to a pure bending moment in the manner shown in Fig. 41 and the geometry of its upper surface determined by laying an optically-flat piece of glass on it and illuminating it vertically with parallel light. The system of interference fringes observed forms two families of hyperbolae (Fig. 42) which may be regarded as *contours* drawn through these points on the top of the specimen which are at equal distances from the optically-flat plate. The ratio of the radii of curvature along and across the bar is equal to the inverse ratio of the spacing of the hyperbolae in these two directions.

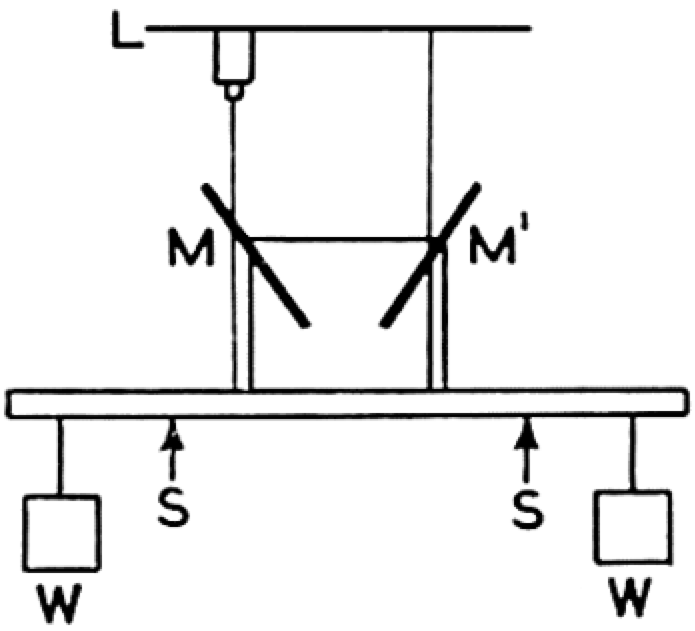


FIG. 41.

(*e*) *Measurements of n.*—If the specimen is in the form of a thin cylindrical rod, the best method is the static one due to **Barton.** One end of the specimen is clamped, and a short cylindrical rod fixed to the other. A pure couple is applied to this end by means of threads wound round the rod, the threads being wound in opposite directions, then passing over pulleys and carrying scale-pans, into which equal weights can be placed. The twist can be measured by attaching a pointer to the rod, but this is very crude, and much better results can be obtained by the use of a galvanometer mirror and scale, which enables twists of a small fraction of a degree to be measured accurately.

A dynamical measurement of n can be made by replacing the weights and pulleys by a rectangular bar, whose moment of inertia can be determined by weighing it and measuring its dimensions, the period of torsional oscillations of the system being determined by timing it. Then, by equation (6) (p. 87), if I is the moment of inertia of the bar and it is twisted through an angle ϕ, the couple acting on it is $\frac{n\pi a^4}{2l}\phi$, and the equation of motion is

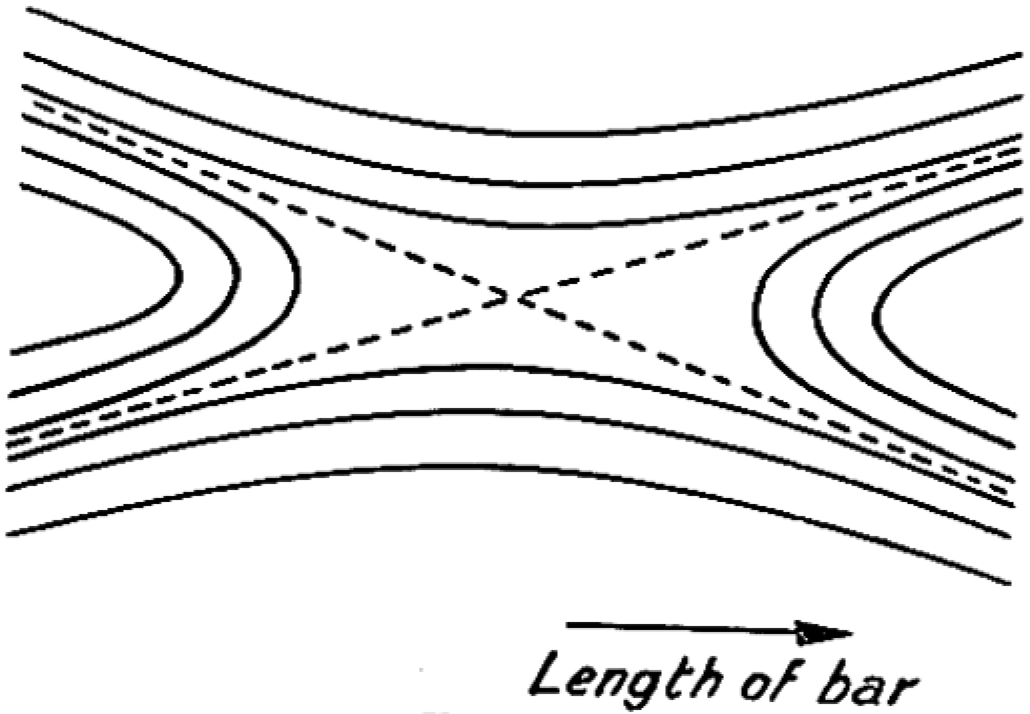


FIG. 42.

$$I\ddot{\phi} + \frac{n\pi a^4}{2l}\phi = 0$$

so that the period of the oscillation is

$$T = 2\pi\sqrt{\frac{2lI}{n\pi a^4}}.$$

(*f*) *Measurement of E and n for the same specimen.*—For a measurement of n, a cylindrical specimen is almost essential, on account of the complexity of the mathematics in other cases, but such a shape is not very convenient for a bending experiment and, for a stretching experiment, it is not easy to choose dimensions that give reasonable values for both extension and torsion with loads that can be conveniently applied in the laboratory. However, equations (18) and (19) (pp. 94-5) show that, if the specimen is wound into a flat spiral spring, E and n can be determined respectively by the torsional and up and down oscillations of the spring if bodies of suitable known moment of inertia and mass are attached to it.

If the specimen is in the form of a short thick wire, a method due to G. F. C. Searle is very convenient. The apparatus shown in Fig. 43 is used, the ends of the specimen being threaded and screwed into equal rectangular bars of square section. To measure E, the apparatus is hung up by strings at points A and B, C and D are drawn together by a thread so as to bend the specimen into an arc of a circle, and the apparatus is set into oscillation by burning the thread. If the length of the wire is l and the angle between the bars 2α, the radius of curvature of the circular arc into which the wire is bent is $l/2\alpha$, and the bending moment is therefore equal to $2EJ\alpha/l$ from equation (11) (p. 90) which is the restoring couple tending to reduce α when the system is released, the same couple acting on each bar. The periodic time of the oscillation is—

$$t_1 = 2\pi\sqrt{\frac{lI}{2EJ}}$$

where I is the moment of inertia of each bar. For a circular section J is equal to $\frac{\pi}{4}a^4$, where a is the radius of the wire, so that

$$t_1 = 2\pi\sqrt{\frac{2lI}{\pi E a^4}}.$$

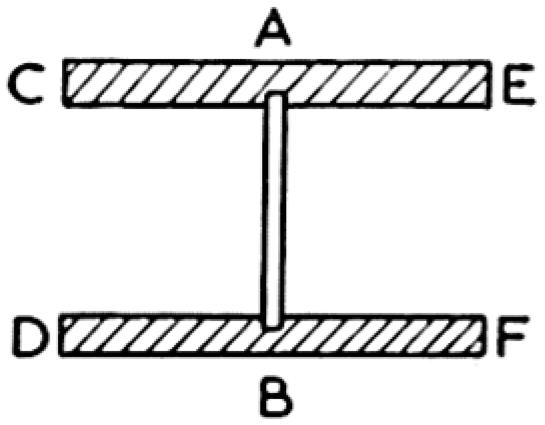


FIG. 43.

If now the bar CAE is clamped horizontally with the wire vertical, and DBF is made to execute torsional oscillations about B, the period of oscillation is given [section (e) above] by

$$t_2 = 2\pi\sqrt{\frac{2lI}{n\pi a^4}}.$$

The relevant value of I is the same in both experiments if the bars are made of square or circular cross-section. We then have

$$\frac{E}{n} = \frac{t_2^2}{t_1^2}.$$

(g) *Measurement of the bulk modulus K.*—A direct measurement of K can be made by means of **Bridgman's** apparatus, described briefly above (p. 78), but simpler methods are also available. For example, if the same pressure is applied to both the inside and outside of an *isotropic* hollow vessel, its volume does not stay constant, but both the internal and external volumes contract in the same ratio as would a solid block of material. Again, if a cylindrical tube, whose radius is large compared with its thickness, is subjected to an internal pressure, it was shown by **Mallock** that the *elongation* depends solely on the bulk modulus, and for an isotropic substance

$$\frac{\delta l}{l} = \frac{PR_1}{6K\,(R_1 - R_2)},$$

where R_1 and R_2 are the internal and external radii. A measurement of the elongation under a known pressure thus enables K to be measured.

7. Plastic flow

As has already been indicated above (Art. 2), the actual stress in a material such as steel does not change very much once the yield-point has been passed. As we have mentioned, only in certain cases may we design structures so that any member is stressed into the plastic region, but the existence of this region is nevertheless of very great engineering significance, for it makes such operations as riveting, wire-drawing, rolling, and pressing possible. Of recent years a great deal of attention has been paid to the investigation of what actually occurs in this region, and a brief account of some of the important conclusions will now be given.

In the first place, experiments with specimens made out of single crystals of metal have shown that they are quite surprisingly *weak*. A photomicrograph of an "ordinary" piece of metal shows that it consists of a multitude of very tiny crystals, their size depending upon the history. "Annealing" a metal amounts to heating it to a temperature at which the crystals can grow by coalescing with one another and then cooling it slowly so as to avoid internal stresses. "Cold working" a metal consists of taking it beyond its elastic limit at a temperature so low that the crystals, once broken up, only join together again very slowly. Now it is well known that annealing a metal softens it and diminishes its strength, while cold working, within limits, does the exact opposite. Part of the explanation of the fact that metals like annealed copper have no definite yield-point is that the breaking up of the individual crystals hardens the metal, and this process continues throughout the deformation, so that the metal becomes progressively harder. A similar process often occurs with steel (p. 77).

We are thus faced with the paradoxical situation that crystal boundaries in a metal are a source of strength, not of weakness. The deformation of a single crystal of metal takes place at very low stresses, and has been found to consist of slipping along certain crystallographic planes known as *glide planes*. Now it obviously requires a smaller stress to slide one layer of atoms parallel to another than it would to tear the two layers apart, but the cohesive forces in metals can be predicted, in order of magnitude, from the electron theory of metals, and it is quite incredible that a *perfect* crystal could pass from the configuration (*a*) to the configuration (*b*) (Fig. 44), under the very small stresses (sometimes no more than a few pounds per square inch) that are required for plastic glide in actual metals.

The accepted explanation is that metallic crystals, however carefully prepared, are never perfect, but contain defects in the lattice structure known as *dislocations.* These may be of many kinds, such as the absence of an atom from a lattice-site or the absence of part of a row or layer of atoms. If a single atom were absent, plastic flow could take place by a *displacement of the vacancy one space at a time*, a process that requires very much less energy than the simultaneous sliding of a whole row. We are thus led to think of plastic glide as a *displacement of dislocations.* In a small crystal the dislocations can only move a short distance, as they cannot run beyond a crystal boundary, whereas in a large crystal with the same number of dislocations per unit volume as the small one, a larger relative deformation is possible as a given dislocation can, on the average, "run" much further than in the small crystal. One thinks of the dislocations as being continually formed in the crystal by thermal agitation and then running as far as possible through the crystal in directions determined by the applied shearing stress. It is possible to prevent the dislocations from running right through a crystal by alloying with the metal in question another metal, whose atoms are sufficiently different in size to form inclusions that distort the regular structure of the crystal, a dislocation being prevented from running further if it reaches such an inclusion. One process of producing inclusions with such properties is known as *precipitation hardening,* because one begins with a homogeneous melt and the inclusions appear only as the alloy cools down.

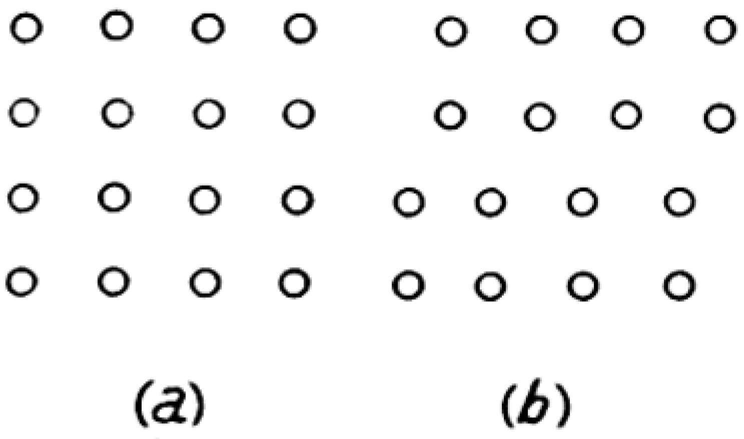


FIG. 44.

8. Variation of elastic constants with temperature

As one might perhaps expect, elastic moduli usually decrease as the temperature rises, but such behaviour is not absolutely universal and some definite exceptions are known. Various empirical expressions have been used for the variation with temperature of the elastic constants, but no definite theory of these variations seems to exist. One question that does arise is whether there is any reason to expect any appreciable difference between elasticities measured under isothermal and adiabatic conditions. Such an effect is well known to be important in the propagation of sound in gases and also in surface tension effects (see p. 233). Calculations, using the observed variations of the elastic constants with temperature, can in fact be made thermodynamically, on the same lines as the calculation of the difference between surface tension and surface energy (p. 233) but the expected difference between the isothermal and adiabatic values usually amounts to only a few per cent. This is only just outside the errors of measurement and the possible range of variation of the constants between different specimens. It will be noticed that nearly all the formulae derived in Art. 6 above involve high powers of some of the dimensions of the specimen so that any errors in measuring these, or any non-uniformity in them, have a disproportionately large effect. One might try to detect the difference between adiabatic and isothermal elasticities by measuring the static and dynamic values for the same specimen. These values are usually not quite alike, but it is uncertain how much of the difference is to be attributed to this effect and how much to other causes such as friction between the small crystals of metal.

As one might expect, increase of temperature greatly facilitates *plastic* flow, and, for this reason, many machining operations such as rolling and pressing, are often carried out on hot metal. A rise in temperature both increases the number of dislocations, and makes it easier for them to "run".

9. Elastic stability

Up till now we have considered the problem of determining the deformation of a body under stress, and we have implicitly assumed that the deformed state is stable until part of the body is stressed into the plastic region. There are, however, certain types of case in which this treatment is inadequate, the most important being that of a beam subjected to a longitudinal *compressive* load, that is to say, a *strut*. It is common knowledge that a strut fails by *buckling*, which is entirely different from the plastic flow and "necking" that heralds the failure of a wire or rod under tension.

The criterion for the buckling of a strut was first obtained by **Euler.** Consider a uniform light rod subjected to an axial compressive load P. The straight configuration is an equilibrium one for *any* load, but we have to decide whether this equilibrium is stable or not. When P is gradually increased to the buckling value, the observation is that the rod first of all begins to bend, but still holds up the load, and that complete failure occurs if the load is increased a little beyond this. We suppose that the original direction of the rod is the y-axis of coordinates and that it is deformed by the load into the curve $x = p(y)$. In Fig. 45

we have two representative cases, case (*a*) representing the situation when the two ends of the strut are constrained to remain in the same vertical line, but are not "built-in", while case (*b*) represents the case where the bottom of the strut is built into the ground, but the top and bottom ends need not be in the same vertical line.

In Fig. 45(*a*) there is no external couple acting on the strut because there is nothing to prevent the lower end changing its inclination relative to the ground, as we have only postulated that its *position* is fixed. From the equilibrium of the portion of the strut above the point Q, whose coordinates are (x, y) we have, by equating the moment of the force P about Q to the bending moment at Q,

$$M_Q = Px.$$

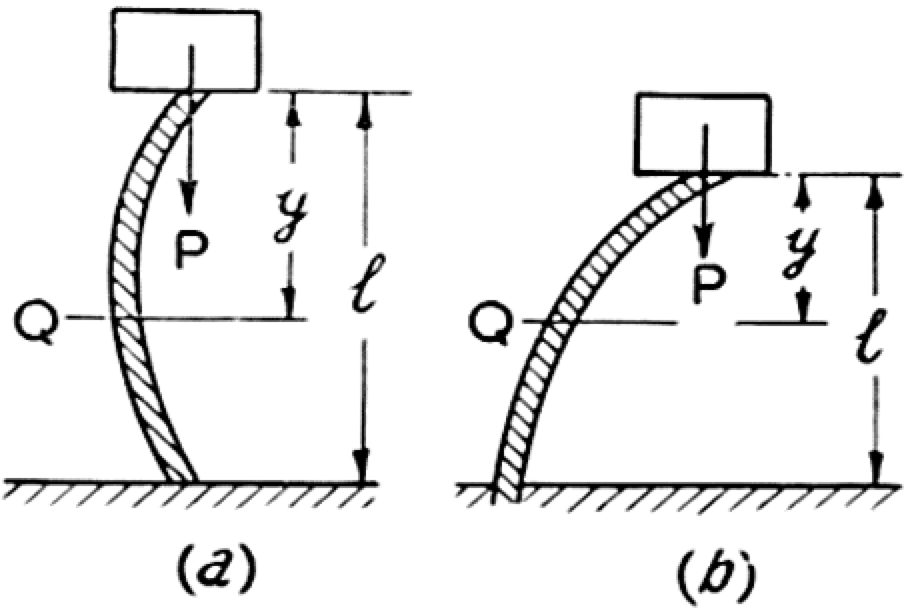


FIG. 45.

Since the strut is supposed deformed into the curve $x = p(y)$, the bending moment at Q is given (p. 90) by the expression

$$M_Q = -EJ\frac{d^2p}{dy^2}$$

and we arrive at the following differential equation for the function p,

$$\frac{d^2p}{dy^2} + \frac{P}{EJ}p = 0 \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (20)$$

Exactly the same equation will hold for the case illustrated in Fig. 45(*b*) provided that we choose the origin so that $p = 0$ at the free end. The fact that the free end is not vertically above the fixed end means that a *couple* is being exerted on the whole strut by the "built-in" joint with the ground. [In Fig. 45(*a*) this is not possible, because the strut is free to pivot in the ground, and for that reason it is necessary to postulate that its two ends are kept in the same vertical line by some external constraint, which would only involve very small forces.]

The general solution of equation (20) is

$$p = A \cos ky + B \sin ky,$$

where $k^2 = P/EJ$. We are still free to choose the zero for the y coordinate and we choose it so that $y = 0$ at the free end. Since p is to vanish here we must have $A = 0$. k now has to be chosen so that the boundary conditions at the ground are satisfied. We let l be the vertical distance of the free end above the ground, which will be very nearly equal to l_0, the undeformed length of the strut, unless the deflections are large. In Fig. 45(*a*), the boundary condition is that $p = 0$ at $y = -l$, since the two ends are in the same vertical line. This is only possible if $\sin kl = 0$, which, if we suppose P increased gradually from zero, becomes possible for the first time when $kl = \pi$, so that the smallest value of P for which equation (20) can be satisfied is

$$P = \frac{EJ\pi^2}{l^2}.$$

This is known as *Euler's criterion for buckling*. In Fig. 45(*b*) the boundary condition at $y = -l$ is not $p = 0$, because this would mean that the top and bottom ends were in the same vertical line, but $dp/dy = 0$ at $y = -l$, since the bottom end is clamped, and must therefore be vertical. dp/dy is proportional to $\cos ky$, and the smallest value of P for which $\cos kl$ can be zero is given by $kl = \pi/2$, that is

$$P = \frac{EJ\pi^2}{4l^2}.$$

Other cases can be treated in the same way. If, for example, in Fig. 45(*b*) we postulated that the strut is to be "built-in" to the load as well as into the ground, so that the strut has a vertical slope at both top and bottom ends, we might at first sight think that we have too many conditions to be satisfied at the ends. We must, however, remember that a "built-in" joint in general implies a couple at the end, and equation (20) ceases to be appropriate if the origin of x is taken at the top end because, instead of the bending moment at Q being equal to the couple due to the external load P, it is equal to this couple plus an unknown constant couple acting at the built-in joint, so that an unknown constant now replaces zero on the right-hand side of equation (20). This, however, can be allowed for by changing the origin from which x and p are measured, so that the top end no longer corresponds to $p = 0$. The appropriate solution of equation (20) is therefore now

$$p = A \cos ky + B \sin ky,$$

where dp/dy is to vanish at $y = 0$ and $y = l$. The criteria for this are

$B = 0$; $kl = \pi$, which leads to the same condition as for Fig. 45(a),

$$P = \frac{EJ\pi^2}{l^2}.$$

The physical meaning of this criterion of Euler's is that the straight configuration is stable for all values of P less than the critical one, but that, for this value of P, the distorted configuration is a possible equilibrium one, however large the multiplying constants A or B may be, so that a slight further increase in P should lead to the deflection increasing indefinitely. This criterion is in good agreement with practice. For a more accurate discussion of the validity of Euler's criterion the reader is referred to Southwell's *Elasticity*.

Another problem of the same kind is the *whirling of shafts*. This occurs if a shaft is rotated at too high a speed to be safe, whereupon its axis deforms into a curve, and it "flops" around in this form, but straightens again if the speed is reduced. Suppose that the shaft is rotating with an angular velocity ω, and is deformed into the curve $y = p(x)$, x being measured along the axis of rotation. Then, if M is the mass per unit length, the centrifugal force, acting on unit length given by $My\omega^2$ and the profile of the shaft, is the same as it would be under an equal static loading. By equation (12) (p. 90) the relationship between loading and deflection is

$$w = -EJ\frac{d^4p}{dx^4},$$

so that possible forms of the function $p(x)$ must satisfy the equation

$$\frac{d^4p}{dx^4} - \frac{M\omega^2}{EJ}p = 0 \qquad (21)$$

the solution of which is

$$p = A\cosh\lambda x + B\sinh\lambda x + C\cos\lambda x + D\sin\lambda x$$

where

$$\lambda^4 = \frac{M\omega^2}{EJ}.$$

We shall suppose that the shaft passes through bearings at $x = 0$ and $x = l$, but consider them to be very short, so that the shaft may stick through them in any direction. There is thus no couple at the ends of the shaft, and, at these points, we have $d^2p/dx^2 = 0$ as well as $p = 0$. We thus obtain four equations between A, B, C, D, and λ, and they can only be satisfied if $A = B = C = 0$, $\sin\lambda l = 0$, the smallest value of ω for which this is possible corresponding to $\lambda l = \pi$, or

$$\omega^2 = \frac{EJ\pi^4}{Ml^4}.$$

For *very long* bearings, the appropriate boundary conditions are $p = 0$, $dp/dx = 0$, but the corresponding equations between A, B, C, D, λ, are then harder to solve.

There exist infinitely many higher values of ω, corresponding to $\lambda l = n\pi$, where n is an integer, just as there are, in theory, infinitely many "buckling loads" for a strut, but, in practice, only the lowest critical whirling speed is of physical interest.

10. Elastic waves

If a system of stresses is suddenly altered there is a corresponding motion of the medium, and a propagation of changes of stress from place to place, the two phenomena occurring together. The whole process is referred to as the propagation of elastic or stress waves. Since, even in an isotropic medium, a deformation in any direction necessarily implies deformation in the two directions at right angles, it follows that the general theory of elastic waves is extremely complicated. In earthquakes the elastic properties of the earth vary with depth, a further complication. In a uniform isotropic medium, we can distinguish three elementary types of elastic wave, and we shall now describe a simple case of each.

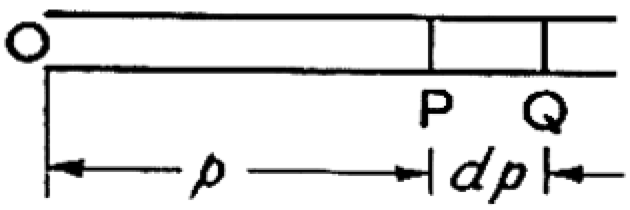


FIG. 46.

(*a*) *Compressional waves.*—These occur, *e.g.*, if a long rod is struck along its axis. Consider (Fig. 46) the motion of the portion PQ of the bar, which we suppose to be of length dx, and cross-section A, and let $p(x)$ be the distance from the origin of a portion of the bar which was originally at x. Then a length of the bar that was originally dx has become dp and the ratio of the new length to the old length is $\partial p/\partial x$, and the fractional extension is $\frac{\partial p}{\partial x} - 1$. The force exerted on the element PQ by the portion of the bar to the left of it is then $-EA\left(\frac{\partial p}{\partial x} - 1\right)_P$, while the force exerted by the portion of the bar to the right of PQ is similarly $EA\left(\frac{\partial p}{\partial x} - 1\right)_Q$. These two do not quite balance, because $\partial p/\partial x$ may have changed slightly between P and Q, and any net force on the element will accelerate it. Thus we have, net force on PQ

$$= \text{mass} \times \text{acceleration}$$

$$= \rho A dx \frac{\partial^2 p}{\partial t^2}$$

(since $\partial p/\partial t$ is the velocity of PQ)

$$= EA\left[\left(\frac{\partial p}{\partial x}\right)_Q - \left(\frac{\partial p}{\partial x}\right)_P\right]$$

$$= EA dx\left[\frac{\partial^2 p}{\partial x^2}\right].$$

Thus the equation of motion is

$$\frac{\partial^2 p}{\partial t^2} = \frac{E}{\rho}\,\frac{\partial^2 p}{\partial x^2}.$$

As is shown in the Appendix (p. 296), this is the equation of waves propagated with a velocity $\sqrt{E/\rho}$. We have tacitly assumed, by using Young's modulus E, that the sides of the bar are left free to move. If we supposed that they were held fixed, the appropriate elastic modulus is not E, but is instead

$$\frac{E\,(1-\sigma)}{(1+\sigma)\,(1-2\sigma)}$$

where σ is Poisson's ratio. This may be written in various other ways, the simplest of which is $K + 4n/3$. (For a liquid or gas n is zero, and the phenomenon reduces to the familiar propagation of sound waves in a tube.) There are thus various types of compressional waves, but they all have two features in common. *The motion is longitudinal, that is, along the direction of propagation of the wave, and it is propagated with a velocity whose square is the ratio of the appropriate modulus of elasticity to the density.*

(*b*) *Shear waves.*—The simplest example of a shear wave occurs as a result of the torsion of a circular bar. Suppose $\partial\theta/\partial x$, the twist per unit length is not constant, as it is in the static case, but varies along the length of the bar in some other manner. The couple associated with this twist is still obtained from equation (6) (p. 87) by replacing ϕ/l by $\partial\theta/\partial x$, so that

$$C = \frac{n\pi a^4}{2}\,\frac{\partial\theta}{\partial x}.$$

This couple varies as we travel along the bar and, acting on the element lying between the planes x and $x + dx$, is a net couple

$$\frac{n\pi a^4}{2}\left[\left(\frac{\partial\theta}{\partial x}\right)_{x+dx} - \left(\frac{\partial\theta}{\partial x}\right)_x\right] = \frac{n\pi a^4}{2}\,\frac{\partial^2\theta}{\partial x^2}\,dx.$$

The (mechanical) moment of inertia of this element is $Ma^2/2$ by Routh's rule, and we have

$$M = \rho\pi a^2 dx,$$

so that the mechanical moment of inertia is $\frac{\rho\pi a^4}{2}dx$, and the angular acceleration is given by

$$\frac{\partial^2\theta}{\partial t^2} = \frac{n}{\rho}\frac{\partial^2\theta}{\partial x^2}.$$

This is a result of just the same type as for compressional waves, but the appropriate elastic modulus is now n. In contrast to compressional waves, shear waves are wholly transverse, the motion being perpendicular to the direction of propagation.

(c) *Flexural waves.*—As an example of these, consider what happens if the loading on a uniform beam is suddenly changed. In Fig. 47, let an element PQ of the deformed beam be of length dx, and let $S(x)$ be the shearing force resulting from the deformation. Then the net force accelerating the element dx upwards is $dx\frac{\partial S}{\partial x}$, which, from equations (10) and (11) (p. 90) give us

$$dx\frac{\partial S}{\partial x} = - EJ\,\frac{\partial^4 p}{dx^4}\,\partial x,$$

where $y = p(x)$ represents the profile of the distorted beam, and equating this to the force needed to give the element an acceleration $\partial^2 p/\partial t^2$ upwards, we obtain

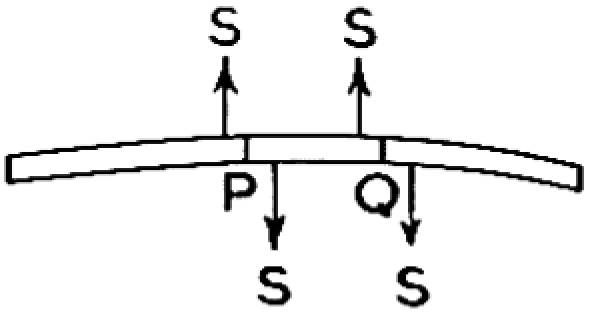


FIG. 47.

$$Mdx\,\frac{\partial^2 p}{\partial t^2} = - EJdx\,\frac{\partial^4 p}{\partial x^4}.$$

[In deriving equations (10) and (11) a deflection p in the downward direction was reckoned positive.] Thus the equation governing displacements involving bending is

$$\frac{\partial^2 p}{\partial t^2} + \frac{EJ}{M}\,\frac{\partial^4 p}{\partial x^4} = 0 \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (22)$$

This equation is of a more complicated type than the simple wave equation. A discussion of its solutions is beyond the scope of this book, but we shall show here that it has the characteristic property that the velocity of propagation *depends on frequency.* For if we try to satisfy equation (22) by a travelling sinusoidal wave

$$p = p_0 \cos\omega\left(t + \frac{x}{u}\right)$$

we find that we can satisfy this equation if

$$u^4 = \frac{EJ}{M}\omega^2,$$

that is the velocity of propagation *increases with frequency.* This means that, unless they are sinusoidal, flexural distortions *change their form* as they travel along, and do not travel forwards or backwards unchanged as do, *e.g.*, disturbances on a stretched string. In this respect flexural waves are analogous to light waves in a material medium, the refractive index of which varies with frequency. (For very high frequencies the theory breaks down, because the wave-lengths and corresponding radii of curvature become comparable with the thickness of the beam.)

Equation (22) can also be used to determine the characteristic frequencies of transverse vibration of beams. If the beam is vibrating at a frequency ω, we may assume a solution of equation (22) of the type $p = q(x) \sin \omega t$, analogous to the stationary waves in an organ-pipe or on a stretched string. Substitution of this expression in equation (22) leads to the equation

$$\frac{EJ}{M\omega^2} \frac{\partial^4 q}{\partial x^4} - q = 0$$

for the function $q(x)$ determining the shape into which the beam deforms. This is of the same type as equation (21) governing the whirling of shafts. If the two ends of the beam are simply supported at $x = 0$ and $x = l$, the differential equation and the boundary conditions

$$q = 0 \text{ (no deflection at ends),}$$

$$\frac{\partial^2 q}{\partial x^2} = 0 \text{ (no couple at ends)}$$

can be satisfied by the simple sine functions

$$q = q_0 \sin \frac{n\pi x}{l},$$

which gives

$$\omega^2 = \frac{n^4 \pi^4 EJ}{Ml^2}$$

as the possible values of ω. This means that the possible frequencies of vibration of a simply-supported beam are proportional to the squares of the natural numbers, and are thus all integral multiples of the fundamental frequency, and are therefore acceptable musically. For a rod or beam with one or both ends clamped $q(x)$ involves hyperbolic as well as trigonometrical functions, and the possible frequencies are not rational multiples of one another, but are solutions of a transcendental equation. For this reason no serious musical instrument uses *clamped* rods, though the xylophone uses *supported* rods or plates.

The main importance of the study of elastic waves is in the design of structures. Suppose, for example, that a beam supported at the ends is given a transverse blow in the middle. The motion is transmitted from point to point by flexural waves, so that the impulse will be transmitted to the supports in a time comparable with that of the passage of a flexural wave from centre to support, and the centre is brought to rest in a time comparable with that of the passage of the wave to the supports plus the time taken by the waves reflected at the supports to travel back to the centre. Roughly, one may say that the centre of the beam does not "know" that the ends are fixed until a wave has had time to make the outward and inward journeys. Until then, the motion of the centre will be the same whether the ends are clamped or free. Another important field of application of the theory of these waves is to the study of *earthquakes,* which represent the energy released by the relative motion of portions of the earth's crust. The earth is not an isotropic medium, and one must take account of the effect of increased pressure leading to more rapid wave-propagation at greater depths. Much interesting information can be deduced. It is, for example, known that most of the inner core of the earth is incapable of transmitting shear waves, which implies that it must be liquid or gaseous.

BIBLIOGRAPHY FOR CHAPTER IV

E. A. Love, *A Treatise on the Mathematical Theory of Elasticity.* Dover Publications, New York (1944).

R. V. Southwell, *Introduction to the Theory of Elasticity for Engineers and Physicists.* 2nd Edn. O.U.P. (1941).

P. W. Bridgman, *The Physics of High Pressures.* 2nd Edn. Bell, London (1949).

H. Jeffreys, *The Earth.* 2nd Edn. C.U.P. (1929).

P. P. Ewald, T. Pöschl, and L. Prandtl, *The Physics of Solids and Fluids.* Translated J. Dougall and W. M. Jeans. 2nd Edn. Blackie (1936).

A. Nádai, *Plasticity* (Engineering Society's Publications) Maple Press, York, Pa., U.S.A. (1931); 2nd Edn. published as *Theory of Flow and Fracture of Solids,* Vol. I. McGraw-Hill, New York (1950).

Examples on Chapter IV

1. Explain why increases of stress may be expected in the neighbourhood of a hole in a stressed member.

2. Derive the result quoted on p. 108 for the modulus of elasticity governing the stretching of a bar whose sides are prevented from moving. (Hint.—Imagine a small cubical element subject to tensions P, Q, Q, along the x, y, and z axes.

The fractional deformation due to P is P/E along the x-axis and $-\sigma P/E$ along the y and z axes. Write down similar expressions for the deformations due to the other two tensions, then adjust the ratio of P to Q so that there is no deformation in the y or z directions.)

3. Apply the method of dimensions (Chap. X) to determine the manner in which the couple exerted by a twisted wire varies with its radius, given that it is proportional to the rigidity modulus and the angle of twist per unit length. (Hint.—Remember that the elastic moduli are all forces per unit area.)

4. By direct integration of equation (16), p. 93, show that a portion of a uniform beam subjected to a constant bending moment forms an arc of a circle. How can this be realised in practice?

5. Obtain the relation between the length and radius of a circular bar which will stand the same load in compression or tension without permanent deformation. [Assume it loaded in compression in the manner shown in Fig. 45(*a*) and that the elastic limit corresponds to a fractional deformation of 0·005.]

6. It is proposed to measure the elastic constants of a wire by subjecting it to various loads, and measuring simultaneously the elongation of the wire and the electrical resistance of the stretched portion.

On the assumption that the specific resistance (resistance of a centimetre cube of the material) is unaffected by stretching, derive a relation for Poisson's ratio for the wire in terms of observed quantities.

7. Calculate the velocity of propagation of sound

(*a*) in a thin steel bar;

(*b*) in a large block of steel.

($E = 2 \times 10^{12}$ dynes/cm.^{2}; $\sigma = 0{\cdot}3$. Density = 8 g./c.c.)

Would you expect to find any differences between the velocities of propagation of sound in a narrow tube and in a large volume for

(*a*) a Newtonian liquid;

(*b*) a gas?

Give reasons.

8. Describe two methods of measuring Poisson's ratio. In what circumstances would you expect a direct measurement of this quantity to give a result differing from that calculated from Young's modulus E, and the modulus of rigidity n.

9. A Kater pendulum has a length from point to centre of heavy bob of 20 cm., an average cross-sectional area of 0·1 cm.^{2}, and a bob of weight $\frac{1}{2}$ kg. If its Young's modulus is 10^{12} dynes/cm.^{2}, calculate the fractional increase in the period due to the elongation of the pendulum by centrifugal force when it is swinging with an amplitude of 1 degree. (Hint.—It is sufficiently accurate to consider the elongation caused by the *mean* angular velocity of the pendulum.)

10. Give a qualitative account of the motion of an anvil when struck at its centre by a hammer. Treat the ends of the anvil as rigidly fixed. In what ways does the motion alter as the direction of the blow is changed?

CHAPTER V

THE KINETIC THEORY OF GASES

1. Fundamental assumptions

In Chapter IV we have discussed the properties of solids with only an occasional reference to their molecular structure, it being satisfactory, for many purposes, to treat them as *continuous*. This point of view is also fruitful for liquids, many of the properties of which can be interpreted by hydrodynamical methods (Chapter VI), which again neglect the molecular structure. On the other hand, the kinetic theory of matter has had its greatest successes for gases, because, in a gas, the molecules are so far apart that the interactions between them are fairly small. For many purposes it is satisfactory to treat the molecules of a gas as *attracting rigid spheres*, that is to say, one considers the molecules of a gas to be in ceaseless motion and continually "colliding" with one another, two molecules being unable to approach closer than a certain distance, usually called the *effective molecular diameter*. This approximation might seem to be exceedingly crude except for monatomic molecules, because most other molecules are by no means spherical. However, diatomic molecules almost certainly *rotate* about axes perpendicular to their lines of centres at all but the very lowest temperatures so that they present a roughly spherical *target* to another molecule approaching them. Another effect that ought to be considered in an accurate theory is the fact that the distance of closest approach will depend on how vigorously the atoms collide, which in turn depends on the temperature. There is some evidence for the existence of such an effect, particularly for the larger molecules such as carbon dioxide.

Even with the simplifying assumptions just mentioned, an *exact* calculation of such a quantity as the viscosity of a gas is still a hopeless task. However, the kinetic theory is able to make very satisfactory predictions for gases, and we must examine why this is so. The main reason is that a gas at ordinary pressures and temperatures consists very largely of empty space, the effective volume of the molecules being a very small fraction of the volume of the gas. Thus, simplifying assumptions can be made that would not be acceptable for liquids and solids, the molecules of which are closely packed. Another important reason is that the molecules are extremely small, so that any volume of gas that we are likely to deal with contains a very large number of them, and the theory of probability can be applied quite safely to determine the *average* values of such quantities as the speed of molecules over a finite area or volume. Indeed, most experimental measurements on a gas can be regarded as determinations of various average properties of the molecules.

2. Maxwell's distribution law for molecular velocities

If we consider a gas "at rest" in an enclosure, we know that its individual molecules are by no means at rest. What we do mean is that there is no "preferred" direction in the enclosure, so that, for every molecule that has a certain velocity component in the x-direction, there is, on the average, just one that has an x-component of velocity of opposite sign and practically equal magnitude. Again, a collision in a gas "at rest" will result as often in a decrease as in an increase of the x-component of velocity of a particular molecule. We are thus led to the idea of these velocity components being *distributed in a random manner about a mean value of zero.* This is just the type of situation envisaged by **Gauss** in his classical work on the theory of probability. He showed that any quantity f that was subject to a multitude of small *random* variations would be distributed about its mean value f_m according to the law

$$P(f_1, f_1 + df) \propto e^{-(f_1 - f_m)^2/K^2} df \quad \ldots\ldots\ldots\ldots \quad (1)$$

where $P(f_1, f_1 + df)$ is the probability of one observation giving a value of f between f_1 and $f_1 + df$ and K is a constant. This is known as the **Gaussian error law** and it is much used in the analysis of experiments where the errors are *random* in the sense of being capable of acting *independently* and *in either direction,* as opposed to *systematic* errors. For a proof of the statement (1) the reader is referred to textbooks of probability theory. **Maxwell** realised the possibility that the collisions of molecules in a gas might be considered to be random events in this sense.

The fundamental assumptions made by **Maxwell** were then as follows:—

(*a*) That, in a gas at rest, the three velocity components are each distributed according to a Gaussian error law, there being *no correlation between* (*say*) *the x and y components of velocity of a specified molecule.*

(*b*) That, in a gas at rest, the distribution parameter K, defined by equation (1), is the same for *all portions of the volume* (apart from the possible influence of external fields of force) and that it is the same for *all three components of velocity.*

(*c*) That the magnitudes of the velocities are so large that any ordinary transport process, such as conduction of heat, can be satisfactorily described by imposing a small *drift velocity* on the Gaussian distribution which is otherwise unaltered.

It must be admitted that no logically rigorous proof of the validity of these assumptions exists. They are, however, intuitively reasonable, and their ultimate justification must rest on a comparison of their consequences with experiment. They are certainly not correct for a liquid or solid, nor would they be valid in a gas but for the fact that the mean distances of the molecules are large compared with the effective ranges of the attractive and repulsive forces.

The parameter K^2 measures the "spread" of, say, the x-components of the velocities about their mean value of zero, that is to say, it is a measure of the mean *magnitude* of the velocities, and we are thus led to expect a connection with the temperature, which measures the vigour of thermal agitation. We shall now determine this connection. Consider a portion of the wall of an enclosure, area A, perpendicular to the x-direction. According to assumption (*b*) the distribution law is the same near the wall as it is everywhere else in the gas. We are assuming that the molecules rebound elastically off the wall and that the forces between wall and molecule are of extremely short range. Let us fix our attention on those molecules whose x-component of velocity is near a certain value v_x. During a time dt, let $An\,(v_x)\,dv_x dt$ molecules with x-components of velocity between v_x and $v_x + dv_x$ hit the area A. The x-component of velocity is simply reversed by an elastic collision, so that the rate of destruction of momentum is $2Amv_x n\,(v_x)\,dv_x$. Now $n\,(v_x)\,dv_x$ is the number of molecules with an x-component of velocity between v_x and $v_x + dv_x$ *crossing unit area in unit time*, and this is equal to $v_x N\,(v_x)\,dv_x$, where $N\,(v_x)\,dv_x$ is the number of molecules in this velocity range *within unit volume at any instant*. By expression (1)

$$N\,(v_x) = N_0 e^{-v_x^2/K^2} \quad \dots\dots\dots\dots\dots\dots \quad (2)$$

where N_0 is determined by the condition that the total number of molecules with all possible values of v_x must be equal to N/V, the number of molecules per unit volume, so that

$$\frac{N}{V} = \int_{-\infty}^{+\infty} N\,(v_x)\,dv_x$$

$$= N_0 \int_{-\infty}^{+\infty} e^{-v_x^2/K^2} dv_x = N_0 K \sqrt{\pi} \quad \dots\dots\dots\dots \quad (3)$$

The total rate of destruction of momentum of molecules of all possible positive values of v_x at the area A is equal to PA, where P is the pressure in the gas, so that, substituting for $n(v_x)$, we have

$$
\begin{aligned}
P &= 2m\int_0^{+\infty} v_x^2 N_0 e^{-v_x^2/K^2} dv_x \\
&= \frac{2mN_0K^3\sqrt{\pi}}{4} \\
&= \frac{NmK^2}{2V} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(4)}
\end{aligned}
$$

For the evaluation of the definite integrals in equations (3) and (4) see Appendix (p. 300).

Equation (4) will agree with the ordinary definition of temperature on the perfect gas scale if

$$RT = PV = \frac{NmK^2}{2},$$

so that

$$K^2 = \frac{2kT}{m},$$

where R is the gas constant for the volume V, and $k = R/N$ is a constant that may be called the gas-constant per molecule or Boltzmann's constant. (Its numerical value is approximately $1{\cdot}37 \times 10^{-16}$ ergs per degree.) It is shown in textbooks on heat that the perfect gas temperature scale is identical with Kelvin's thermodynamic scale of temperatures. We have thus determined the relation between the distribution law (2) and the temperature.

The principle of equipartition of energy. Maxwell's distribution law

Our assumption is that all three components of velocity are distributed according to laws like (2), so that, if $N(v_x, v_y, v_z)$ is proportional to the number of atoms with velocities near v_x, v_y, v_z, we have

$$N(v_x, v_y, v_z) = N_1 e^{-m(v_x^2 + v_y^2 + v_z^2)/2kT} \quad \ldots\ldots\ldots \text{(5)}$$

Supposing now that we are interested not in the *direction* of velocity of a molecule, but only in the *magnitude* of the velocity vector. If we write

$$c^2 = v_x^2 + v_y^2 + v_z^2$$

the process of averaging over all possible values of v_x, v_y, and v_z can be replaced by an average over possible values of the variable c. Mathematically, the two methods of averaging stand in the same relation to one another as do the two possible methods of integrating over the whole of coordinate space, using either Cartesian coordinates or spherical polar coordinates. Thus, since the result of integrating a function $f(r)$ over a *sphere* of radius R is $\int_0^R 4\pi r^2 f(r)\,dr$, we must write the result of averaging over all *directions* as

$$N(c) = N_1 e^{-mc^2/2kT}$$

where N_1 is determined by the condition

$$\begin{aligned} \frac{N}{V} &= 4\pi\int_0^\infty c^2 N(c)\,dc \\ &= 4\pi N_1\int_0^\infty c^2 e^{-mc^2/2kT}\,dc \\ &= \pi^{\frac{3}{2}} N_1\left(\frac{2kT}{m}\right)^{\frac{3}{2}}. \quad \text{See Appendix, p. 301.} \end{aligned}$$

We can cover the whole of ordinary space either by integrating over the x, y, and z coordinates between $-\infty$ and $+\infty$, or by integrating over a large sphere. Similarly, we can provide for all possible *directions* of a velocity of magnitude c or less by integrating over a sphere of radius c. Thus the number of molecules in unit volume having velocities of magnitudes between c and $c + dc$ is given by substituting for N_1 in terms of N.

$$N(c)\,dc = \frac{N}{V}\left(\frac{2\pi kT}{m}\right)^{\frac{3}{2}} e^{-mc^2/2kT}\,c^2 dc \quad \ldots\ldots\ldots\ldots (6)$$

This is known as *Maxwell's distribution law*. Two quantities that are frequently needed are the mean square velocity, that is to say, the average of c^2 over all the molecules, which is given by

$$\begin{aligned} \overline{c^2} &= \frac{V}{N}\int_0^\infty 4\pi c^2 N(c)\,dc = 4\pi\left(\frac{m}{2\pi kT}\right)^{\frac{3}{2}}\int_0^\infty c^4 e^{-mc^2/2kT}\,dc \\ &= \frac{3kT}{m} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (7) \end{aligned}$$

and the mean velocity, which is the average of c over all the molecules, is given by

$$\begin{aligned} \bar{c} &= \frac{V}{N}\int_0^\infty 4\pi c N(c)\,dc = 4\pi\left(\frac{m}{2\pi kT}\right)^{\frac{3}{2}}\int_0^\infty c^3 e^{-mc^2/2kT}\,dc \\ &= 4\pi\left(\frac{m}{2\pi kT}\right)^{\frac{3}{2}}\cdot\frac{1}{2}\cdot\left(\frac{2kT}{m}\right)^2 \\ &= \sqrt{\frac{8}{\pi}}\left(\frac{kT}{m}\right)^{\frac{1}{2}} = \sqrt{\frac{8}{3\pi}}\left(\overline{c^2}\right)^{\frac{1}{2}} \quad \ldots\ldots (8) \end{aligned}$$

Comparison of equations (7) and (8) show that $\overline{c^2}$ is not precisely the square of $\bar{c}$, but exceeds it by the factor $3\pi/8$, which is slightly larger than unity, and is a universal constant independent of V, T, etc. This difference arises because the values of the velocity vector are distributed over all values. A cruder form of kinetic theory would not take account of this distribution of velocities, but would assume that all molecules had velocities of nearly the same magnitude.

This sort of situation is very common in kinetic theory. A simple assumption, such as that just mentioned, that all molecular velocities are nearly equal to $(\overline{c^2})^{\frac{1}{2}}$, will lead to the correct *form* of theoretical formula. The usual consequence of making a more exact assumption is either to confirm the formula obtained by the cruder treatment, or else to alter the numerical constants. As these "rough and ready" methods are always considerably shorter than the more exact ones, they are much used when what is wanted is an insight into the physics of a problem rather than a comparison of numerical predictions with experiment. We shall make much use of them in this chapter.

Boltzmann's expression for the distribution of systems among states of different energy

Equation (7) shows that the mean kinetic energy of a molecule depends only on the temperature, it being equal to $\frac{3}{2}kT$, of which $\frac{1}{2}kT$ is associated with each of the components of velocity v_x, v_y, v_z. Thus two gases at the same temperature will have the same kinetic energy per molecule, and, since two such gases can be mixed without altering the temperature of either, we conclude that *the average kinetic energy is the same for each component of a mixture of two or more gases.* This is the celebrated principle of *equipartition of energy.* By the application of statistical mechanics, to textbooks of which the reader is referred for details, **Boltzmann** was able to generalise this result. His conclusion is that, *if a system (such as an atom or molecule) can occupy more than one state, and forms part of an assembly in dynamic equilibrium, the probability of finding it in a state of total energy U (kinetic plus potential energy) is proportional to $e^{-U/kT}$, where T is the temperature of the assembly of which the system forms a part.* This result is of very great generality, and Maxwell's distribution law is a particular case of it.

Sedimentation equilibrium

As a simple example of the application of Boltzmann's expression, we may consider the problem of *sedimentation*, that is, the balance that is struck between the conflicting influences of gravity and thermal agitation, gravity tending to collect all the atoms at the bottom of the vessel, thermal agitation to disperse them. The simplest example is the variation of the density of the atmosphere with height. The potential energy of a molecule at a height h is mgh, so that, if the temperature of the atmosphere were constant, Boltzmann's expression states that the number of molecules in a small column of gas at a height h varies as $e^{-mgh/kT}$, provided that the average kinetic energy of a molecule is independent of h. It may be wondered why there is no change in the average kinetic energy as the potential energy increases. The complete Boltzmann factor is equal to

$$exp\left[\frac{-m}{kT}\left(gh+\frac{v_x^2+v_y^2+v_z^2}{2}\right)\right]$$

which can be averaged over all possible values of the velocities just as in equations (2)-(8) above. The calculation, *e.g.* of $\bar{c}^2$, depends on the *ratio* of the integrals $\int_0^\infty c^2 N(c)\,dc$ and $\int_0^\infty N(c)\,dc$, which ratio is entirely unaffected by a factor $e^{-mgh/kT}$ in $N(c)$, so $\bar{c}^2$ is independent of h. The factor $e^{-mgh/kT}$ introduces a variation of *density* with height, but the law of distribution of *velocities* depends only on m and T, as we have seen above. This conclusion follows from Boltzmann's expression as a consequence of the kinetic and potential energies of an average single molecule varying independently of one another.

The same factor $e^{-mgh/kT}$ can be deduced without the use of Boltzmann's expression as follows, using simply the elementary kinetic theory of gases: consider a column of gas of unit area at constant temperature, and suppose that it is subjected to no pressure except its own weight. Then, if $N(z)$ represents the number of atoms per unit volume at a height z and $P(z)$ the pressure at a height z, then the force on unit area at height h is equal to the weight of all the molecules directly above that area, that is

$$P(h) = mg\int_h^\infty N(z)\,dz \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (9)$$

But, by applying the gas laws, which, as we have seen above, follow from elementary kinetic theory (p. 116), to unit volume at height h,

$$P(h) = RT = N(h)\,kT,$$

since k is the gas-constant for a single molecule. The relation, (9), between P and N can be written in differential form as

$$mgN(h) = -\frac{dp}{dh}.$$

Using the gas laws to eliminate $N(h)$ and integrating the resulting differential equation of the first order, for P we get

$$\log P = \text{const.} - \frac{mgh}{kT}$$

$$\text{or} \quad P = P_0 e^{-mgh/kT};$$

$$\therefore \quad N = N_0 e^{-mgh/kT} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (10)$$

If we take the mass of an average air molecule as 5×10^{-23} g. this would mean that for uniform temperature the density of the atmosphere would drop to $1/e$ of its value at the earth's surface in a height of eight kilometres, which agrees roughly with observation. (The temperature of the atmosphere is not uniform, but varies with height in a complicated fashion.)

A modification of equation (10) is of great historic interest, as it formed the basis of the first accurate determination, by **Perrin**, of atomic masses. Instead of the atmosphere Perrin used suspensions of gamboge particles, of nearly uniform size, in various liquids, and determined the variation of concentration with depth by counting the number of particles visible in the field of a high-power travelling microscope as its height was varied. (Such a microscope has a very small "depth of focus", and only particles within a very small range of depths are visible at any one setting.) He found an exponential variation with height, analogous with equation (10), allowance being made for the buoyancy of the liquid suspending the particles, the revised equation being

$$N = N_0 exp\left[-\frac{gh}{kT}(\rho - \rho_l)\, V\right] \ldots\ldots\ldots\ldots\ldots \quad (11)$$

where V is the volume of a particle, ρ its density and ρ_l that of the liquid. Since all the other quantities in equation (11) are known, k can be determined, and since the gas constant per gram-molecule is known, the number of molecules in a gram-molecule, Avogadro's number, can also be determined.

This fact that a suspension of microscopic particles behaves essentially like a gas, in spite of their very much greater mass, is very striking confirmation of the principle of equipartition of energy. The velocities associated with particles of sizes visible in the microscope are quite small compared with gas-kinetic velocities, but are, nevertheless, quite easily visible and are sufficient to prevent small particles from "settling out". The ceaseless agitation of minute grains of pollen suspended in water was discovered by **Brown** many years before the kinetic theory of gases existed, and this motion of microscopic particles is known as the *Brownian movement*. A systematic statistical analysis of the successive velocities acquired by a particular particle always shows that its motion is *completely random* in direction, exactly as was postulated by **Maxwell** for gas molecules, and thus provides a direct check on one of his assumptions.

Only a few attempts to determine directly the distribution law for velocities of thermal agitation seem to have been made. For *electrons* emitted from a hot filament, a direct determination can be made by plotting current against voltage in a thermionic diode. (The reader is referred to textbooks of electricity for details of this experiment.) One may get a current flow (of electrons travelling from filament to anode) even if the electric field is tending to urge them the other way. By measuring this thermionic current one can determine directly how many electrons have sufficient velocity towards the anode to overcome this electric field, so that we are determining directly the distribution law for velocities, which seems to be truly Maxwellian.

As for neutral gases, it is possible, in principle, to "sort out" different velocities by devices such as mounting two discs on the same shaft, cutting a slit in each, separated by a certain angle, so that only atoms with velocities in a certain small range can pass through the slits in both discs when they are rotated together at a certain speed. Such experiments have been tried, but disturbing effects are large, and one can conclude only that the law is qualitatively verified.

3. The concept of "mean free path"

Having determined the laws of distribution of velocities in a gas "at rest" and in equilibrium, we now have to examine the processes of *restoration of equilibrium* when a gas is disturbed. There are various types of process that can occur, together or separately. There is *diffusion* which is produced by variations of *density* from point to point, *thermal conduction* and *thermal transpiration* which are produced by variations of *temperature*, and *viscous flow* which is produced by variations of *pressure*. Various other phenomena occur in mixtures of gases as a consequence of the different mean velocities associated with molecules of gas of different masses at the same temperature. For example, if a gradient of temperature exists in a mixture, a gradient of concentration also tends to be established, because of the difference in mean velocities, and this effect, known as *thermal diffusion*, can be used for the separation of isotopes, which cannot be separated by chemical means. All these phenomena are profoundly modified at very low pressures (10^{-3} mm. of mercury or less) and various new phenomena appear (Art. 5, p. 130).

Definitions of mean free path.—Quite a number of definitions of mean free path exist, and, in accurate work, it is necessary to pay attention to the differences between them. *Maxwell's* definition is the natural one of the average distance a particular molecule travels between two successive collisions. *Tait's* definition is the slightly different one of considering an assembly of molecules *at a given instant*, and then taking the average distance each of the molecules travels between that instant and its next collision. The idea of a "collision" becomes somewhat vague as soon as we drop the idea of treating molecules as rigid spheres and admit that they exert long-range forces on one another, because then any two molecules passing near one another affect one another's velocities slightly, and a more "direct" impact only increases this effect numerically, and it consequently becomes difficult to decide whether a "collision" has taken place or not. In practice, the following definition of mean free path is extremely useful. We imagine a small drift velocity in, say, the x-direction imposed on the distribution of random velocities in a gas, and ask how far, on the average, does a molecule have to travel before it has completely lost that drift velocity, which distance we *define* as the mean free path. This is practically equivalent to the distance between collisions, if we suppose the molecules to be rigid spheres, for a collision can leave the drift velocity unaffected, destroy it, or reverse it. We may say, roughly, that one collision is enough, on the average, to destroy any drift velocity, those molecules which preserve it practically intact after a glancing collision being approximately compensated for by those molecules whose drift velocity is *reversed* by a head-on collision.

Distribution of distances between collisions about a mean value.—The numbers of atoms that describe, between collisions, distances significantly larger or shorter than the mean free path are sometimes needed. This distribution law is *not* a Gaussian one, but will now be determined. Suppose, at any instant, we have $n(x)$ atoms that have travelled a distance x and still retain their drift velocities. After each has travelled a further distance dx, a certain number of them, $dn(x)$, will have made collisions and will have lost their drift velocities, which number must be proportional to $n(x)$ and to dx, so that we have

$$dn(x) = -\frac{n(x)\,dx}{l} \quad \dots\dots\dots\dots\dots\dots \quad (12)$$

where l is a constant. Integrating, the number $n(x)$ of molecules which preserve their drift velocities after travelling a distance x is given by

$$n(x) = Ne^{-x/l} \quad \dots\dots\dots\dots\dots\dots \quad (13)$$

the constant of integration being determined by the fact that $n(0)$ must equal N, the total number of molecules present. For many purposes, the ultimate consequences of equation (13) are identical with those of the cruder assumption that every molecule loses its drift velocity after travelling a distance precisely equal to l. If equation (13) is multiplied by x and integrated for all possible values of x from zero to infinity, the average path described by a molecule before losing its drift velocity is just l, which is thus the mean free path, as defined above.

The relationship between mean free path and effective molecular diameter.—If we think of molecules as rigid spheres of diameter σ, the *form* of the relationship between l and σ can be derived very simply. For two molecules "collide" if their centres approach closer than a distance σ. Fixing our attention on one particular molecule we may think of each of the others as presenting to it a *circular target* of radius σ, and it will collide with another molecule if its centre is following a path that intersects the corresponding "target". Alternatively, as our particular molecule moves forward a distance dx, we may think of it as sweeping out a cylinder of volume $\pi\sigma^2 dx$, and a collision will occur if the centre of another molecule is found inside this cylinder. If there are N molecules per unit volume of gas, the probability of the centre of one or more being inside this cylinder is $N\pi\sigma^2 dx$, which is thus the probability of a collision occurring in the region dx of the path of a typical molecule. Comparing this with equation (12), we conclude that

$$l = \frac{1}{N\pi\sigma^2} \qquad (14)$$

Thus, if l is known, σ can be determined. σ is known as the "gas-kinetic diameter", and agrees roughly with molecular dimensions determined by other means (see Table VI, p. 315). l, the mean free path, is usually inferred indirectly from measurements of viscosity or thermal conductivity, but the experiments of **Franck** and **Hertz** on critical potentials of gases provide a means of measuring l directly. A gaseous ion, or electron, if it is moving with a sufficiently high velocity, can ionise other molecules by collisions. If a known electric field is applied the ions are accelerated between collisions, but, on an average, the resulting drift velocity is destroyed at each collision, the maximum drift velocity being equal to $\epsilon El/m\bar{c}$, where ϵ is the charge on the ion and E the applied electric field. This follows from conservation of momentum, for ϵE is the force on the ion, while the time it takes to describe a mean free path is $l/\bar{c}$. If the velocity necessary to cause ionisation by collision can be measured in some other manner, l can be determined from the critical value of the applied field necessary to cause ionisation. The energy necessary to ionise an atom is accurately known from spectroscopic data. Owing to the small mass of an electron, it can impart nearly all its kinetic energy to an atom with which it collides "head-on".

Equation (14) has been derived by a very crude argument but is very nearly correct, various refinements that have been made in it tending to cancel one another's effects. The following are some of them:

(*a*) Taking account of the fact that *both* the "typical molecule" and the "target" are in motion [whereas in deriving equation (14) we assumed the "target" to be at rest], multiplies the expression for l by a factor $\frac{3}{4}$.

(*b*) Taking account of the Maxwellian distribution of velocities modifies the factor $\frac{3}{4}$ to $1/\sqrt{2}$.

(*c*) Taking account of the point mentioned above, the persistence of part of the drift velocity after a typical collision, modifies the factor to $1{\cdot}319/\sqrt{2}$ (**Jeans**).

(*d*) Taking account of the fact that the molecules are not rigid, but exert forces on one another, **Sutherland** arrived at the still more general formula

$$l = \frac{1{\cdot}402}{\sqrt{2}\left(1+\frac{K}{T}\right)} \quad \frac{1}{N\pi\sigma^2} \quad \ldots\ldots\ldots\ldots\ldots\ (15)$$

which is practically the same as formula (14) except for the term K/T, K being a constant for each gas, which measures the effective reduction in the molecular diameter as the collisions increase in violence with rising temperature, that is the "softness" or compressibility of the molecule itself. This effect is negligible in monatomic or diatomic gases, but becomes important for such molecules as carbon dioxide.

4. Relations between mean free path and the transport properties of gas

Viscosity.—Viscosity is defined (see Chapter VI) as the tangential force which comes into play when layers of gas are slipping past one another, and is measured by the tangential force acting over unit cross-section when the gradient of velocity is unity. The viscosity of a gas is very simply explained by the kinetic theory, for let Fig. 48 represent two small volumes of gas which are moving relatively to one another because of a velocity gradient in the y-direction. Now this means that a molecule in the upper small volume will have a slightly greater drift velocity than one in the lower small volume. Now the fact that molecules are continually travelling backwards and forwards between these two volumes tends to increase the momentum of the lower volume and to diminish that of the upper one, and thus to equalise the velocities of the two volumes. The corresponding force is the viscous drag.

Consider a surface of unit area perpendicular to the y-axis. The number of molecules crossing it in unit time is $\frac{1}{6}N\bar{c}$, the factor $\frac{1}{6}$ arising because, if we consider a cube of gas of side unity, any particular molecule can leave this cube by crossing any one of the six bounding squares and $N\bar{c}$ is the total number of molecules leaving unit volume per second. Let v_x be the average drift velocity in this neighbourhood which will vary with y. The drift momentum carried upwards by these molecules is mv_x for each one, on the average, corresponding to a total rate of transport of momentum of $\frac{1}{6}N\bar{c}mv_x$. On the average, a molecule after travelling a distance l, has suffered enough collisions to adjust its drift velocity to the value proper to its new surroundings, and its drift velocity is now $v_x + l\frac{dv_x}{dy}$. Thus the rate at which all the molecules crossing the area from below gain momentum is $\frac{1}{6}N\bar{c}ml\frac{dv_x}{dy}$, and those crossing it from above, similarly lose momentum at the same rate. The total force acting on all the molecules crossing this area is thus $\frac{1}{3}N\bar{c}ml\frac{dv_x}{dy}$ and we arrive at the following theoretical expression for the viscosity

$$\eta = \tfrac{1}{3}N\bar{c}ml \quad \ldots . \quad (16)$$

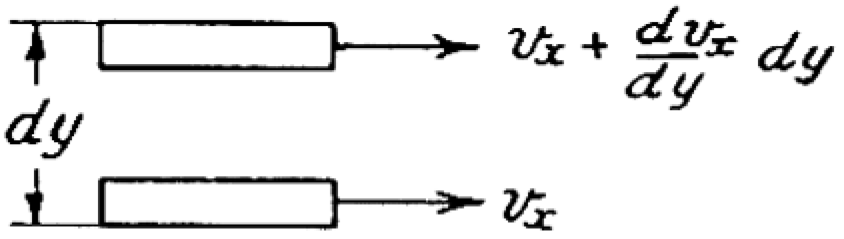


FIG. 48.

It will be noticed that we arrived at this result as a direct consequence of our definition of l as the distance an atom has to travel before its drift velocity is the same as that of its surroundings. This by-passes the tricky questions we have mentioned already, of the number of collisions needed to destroy drift velocity and of what exactly constitutes a "collision" if the molecules are no longer represented as rigid spheres, but it still involves the concealed assumption that the mean free path is the same for molecules of all velocities.

Heat conductivity.—If there is a temperature-gradient in a gas, there is a corresponding variation in the mean velocity $\bar{c}$. If a typical molecule travels from a warm to a cold place, it will lose a little of this velocity and settle down to the mean velocity appropriate to its new surroundings after travelling, on the average, a distance l.

Suppose that there is a temperature gradient along the y-axis, and consider the energy carried across unit area perpendicular to the y-axis. As in the derivation of the viscosity, the number of molecules crossing this area per second from below is $\frac{1}{6}N\bar{c}$. Let E be the *total* energy of a molecule, which will vary with y if there is a temperature gradient. On the average, a molecule, after travelling a distance l, has suffered enough collisions to enable it to acquire the energy proper to its new surroundings, so that the rate at which molecules crossing the area from below gain energy is $\frac{1}{6}N\bar{c} \times$ (energy gained while a typical molecule travels a distance l), that is $\frac{1}{6}N\bar{c}l\dfrac{dE}{dy}$. Similarly, the rate at which molecules crossing the area from above lose energy is the same, and the total rate of transport of energy across unit area is

$$\tfrac{1}{3}N\bar{c}l\frac{\partial E}{\partial T}\frac{dT}{dy} \quad \text{or} \quad K = \tfrac{1}{3}N\bar{c}lmC_v\frac{dT}{dy} \qquad \text{.......... (17)}$$

since $\partial E/\partial T$ is, by definition of specific heat, equal to mC_v, where m is the mass of a molecule and C_v is the specific heat at constant volume. This argument holds whether we have a monatomic gas, in which case the energy E is entirely kinetic, or a polyatomic gas, in which case E also contains a part associated with the internal degrees of freedom of the molecule. Since either of these forms of energy can be lost or gained in collisions, both forms must be taken into account when we are considering the rate of transport of heat. We need not, however, consider the average potential energy in external fields such as gravity, because this is unaffected by collisions (see p. 119).

Comparing formulae (16) and (17) suggests the relation

$$K = \eta C_v \qquad \text{...................... (18)}$$

and this is a relation that can be experimentally tested. More refined arguments, allowing more accurately for the interaction of the molecules, suggest that this simple formula might be in error by a factor of as much as two, so that the results of the comparison are rather unsatisfying, and we can only claim "order of magnitude" agreement between theory and experiment.

Diffusion and allied phenomena

True diffusion is essentially a property of mixtures of gases or solutions. It is not possible to have two portions of a pure gas at different densities without there being also a difference of pressure or of temperature or both, in which case the mechanisms of thermal conductivity and viscous flow will also operate. We might, however, separate two portions of gas at different pressures, but at the same temperature, by a metal foil pierced with small holes, or a filter paper. If the length of the pores is short compared with the mean free path, then the process by which the concentrations are equalised is distinct from viscous flow, and is known as *effusion*. If we have N_1 molecules per unit volume on one side of a pore of area A, the number falling normally on the pore is $\frac{1}{6}N_1\bar{c}A$ (see p. 125), while, if there are N_2 molecules per unit volume on the other side, the number falling normally on the pore from that side is $\frac{1}{6}N_2\bar{c}A$, and, if the length of the pore is short compared with the mean free path, all molecules falling on the pore travel through it, so that

$$\frac{\partial(N_1 - N_2)}{\partial t} = -\tfrac{1}{6}A\bar{c}(N_1 - N_2).$$

Thus for a given piece of apparatus, whose total pore area is fixed, the time-constant associated with effusion is inversely proportional to $\bar{c}$, that is, it is proportional to $\sqrt{T/m}$, since $m\bar{c}^2 = 3kT/2$. This result is known as **Graham's effusion law** and has been used for the determination of molecular weights.

Thermal transpiration.—Let us now suppose that two portions of gas are separated by a porous plate, and maintained at different temperatures. Then, as was discovered by **Osborne Reynolds**, conditions eventually settle down to a state of affairs in which there is no net flow across the plate, yet there is both a pressure and a temperature difference between the two sides. If N_1, N_2, are the number of molecules per c.c. on the two sides, and $\bar{c}_1$, $\bar{c}_2$, the mean velocities, then the condition that there should be no net flow through the plate is evidently that the total number of molecules falling normally on two ends of a typical pore of area A should be equal, that is

$$\frac{N_1\bar{c}_1 mA}{6} = \frac{N_2\bar{c}_2 mA}{6}$$

or, since $p_1 = \frac{1}{3}N_1 m\bar{c}_1^2$ and $p_2 = \frac{1}{3}N_2 m\bar{c}_2^2$,

$$\frac{p_1}{p_2} = \frac{\bar{c}_1}{\bar{c}_2} = \sqrt{\frac{T_1}{T_2}} \qquad \dots\dots\dots\dots\dots\dots \quad (19)$$

Again, this formula holds only if the mean free path is large compared with the dimensions of the pores, that is, at very low pressures. At higher pressures, the effect is disturbed by the backward flow of gas through the pores as the result of ordinary viscous flow, which will begin as soon as any pressure-difference appears. A more refined investigation, covering all pressures, has been given by **West**, who finds that formula (19) is only a limiting one at low pressures, and that at high pressures the ratio of pressure and temperature differences depends on the viscosity of the gas and the size of the pores, while equation (19) suggests a relationship depending only on the temperatures.

Thermal diffusion.—The circumstance that, at higher pressures, the relationship between pressure and temperature differences depends on the nature of the gas has been exploited in the *separation of isotopes.*

If a mixture of two gases of different masses (even if identical chemically) is subjected to the treatment just described, then, at very low pressures, each gas separately will obey equation (19), and there is no difference in the relative concentrations on the two sides of the plate. If, however, a higher pressure is used, the difference in masses of the two types of molecule begins to be effective (principally because of the influence of the molecular mass on viscosity) and, by careful control of the conditions, an appreciable separation can be obtained, even if, as in uranium compounds, the ratio of the molecular masses differs only slightly from unity.

True diffusion.—We have described all these allied effects first in order to draw attention to the difficulty of staging an experiment in a gas in which "true" diffusion can be observed apart from any other phenomenon. The phenomenon is best studied in liquids and solutions, where the influence of pressure is quite small. It is, however, instructive to work out a theory of the diffusion of "foreign" molecules in a vessel containing gas. Pairs of gases have long been known which have practically equal molecular masses and very similar external electronic configurations. From the point of view of the kinetic theory of gases the interactions and effective molecular diameters of such molecules are practically identical, but they can be distinguished chemically. A classic example of such a pair of gases is carbon monoxide (CO) and nitrogen (N_2). Another way in which one can "label" certain molecules is by replacing some or all of their atoms by *radioactive isotopes*, which are chemically identical with ordinary atoms but differ from them in having unstable nuclei, and can thus be used in the same way as "tracer" bullets. Let us suppose, then, that we have a mixture of two gases sufficiently alike to possess a common root-mean-square velocity and mean free path. Suppose that N_a, the number of one type of molecule varies with position along the y-axis, and consider the number of A-molecules crossing the area from below. These have, on the average, travelled a distance l since their last collision, so that this number is

$$\frac{\bar{c}}{6}\left(N_a - l\,\frac{\partial N_a}{\partial y}\right)$$

while the number crossing it from above is similarly

$$\frac{\bar{c}}{6}\left(N_a + l\,\frac{\partial N_a}{\partial y}\right)$$

so that the diffusion coefficient D, defined in a way analogous to that of heat conductivity, as the net mass of gas crossing unit area per second per unit density gradient, is $\frac{1}{3}\bar{c}\,l$. Inserting numerical values for $\bar{c}$ and l, this gives a value of the order of 0·1 C.G.S. units for ordinary conditions. The corresponding figure for the diffusion coefficient of a solid or gas in a *liquid* is very much smaller, 10^{-5} C.G.S. units being a typical value. It is indeed common knowledge that diffusion in liquids and solutions is always comparatively slow, but diffusion in gases is rapid. If, for example, a gas jet is turned on in a room, the smell becomes noticeable at the far side of the room within a few seconds, which is consistent with the diffusion coefficient, in C.G.S. units, being of the order of unity. For the reasons already stated, diffusion coefficients in gases are hard to measure accurately. In Chapter VIII we shall derive the equation governing the variation of concentration of a diffusing solution, but the equation is equally applicable to the diffusion of one gas into another. The kinetic theory leads to the result: rate of transfer $\propto$ concentration gradient, and from this an equation of exactly the same form as that in Chapter VIII (p. 228) can be derived.

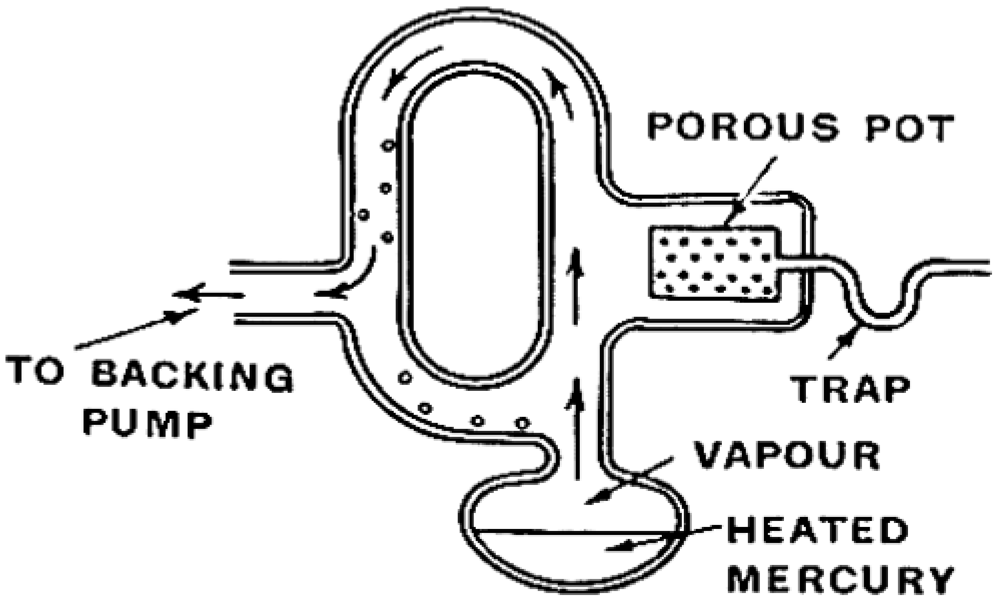


FIG. 49.

Diffusion pumps.—One very important application of the diffusion of gases into one another is the diffusion pump (Fig. 49). The design of such pumps has now become a highly specialised subject, and such pumps are much used when high pumping speeds or extremely low pressures are wanted. The principle of the pumps is, however, extremely simple. Suppose that the vessel to be evacuated is connected to a porous pot, round the outside of which a stream of mercury vapour is flowing rapidly. The gas diffusing out of the pot is rapidly swept away by the mercury stream, so that the pressure of gas inside the pot is constantly falling. Some mercury vapour will no doubt diffuse into the pot, but it can be prevented from reaching the vessel being evacuated by interposing a liquid-air trap (a short length of tube, cooled in liquid-air, which cools down the mercury to a temperature at which its vapour pressure is negligible). When the mercury vapour has left the neighbourhood of the porous pot it can be condensed, and the gas that it has swept away is then removed from the apparatus by a so-called "backing pump" (which can be a mechanical pump as long as it is capable of maintaining the pressure in the apparatus at less than 1 mm. of mercury). The liquid mercury can then be returned to the reservoir to be vaporised once more. In practice, the porous pot we have postulated is replaced by carefully designed slits, the object being to secure maximum outward diffusion of the gas to be pumped and minimum influx of mercury vapour. For extremely low pressures special oils of low vapour pressure have been developed to replace the mercury, and, with some of these oils, the liquid-air traps can be dispensed with.

5. Phenomena at low pressures

One very striking consequence of the kinetic theory of gases is that the coefficients of viscosity and thermal conductivity only involve the mean free path l through the product Nl, and equation (1·4) suggests that this should be a constant, *independent of the pressure.* We thus arrive at the remarkable conclusion, first predicted by **Maxwell**, that *the viscosity and thermal conductivity of a gas should be independent of pressure.* (They are not independent of temperature, because they depend on c.) This very surprising prediction was investigated experimentally and verified down to very low pressures, of the order of a fraction of a millimetre of mercury. (It will be noticed that a prediction of this kind is exempt from the major bugbear of kinetic theory, the uncertainties in the theoretical numerical constants.) From the physical point of view, the result is reasonable enough, the longer mean free path at low pressures just compensating for the smaller number of molecules available to transport momentum or energy. Nevertheless, the quantitative verification of this conclusion must be regarded as one of the major triumphs of kinetic theory, and it is strong evidence for the soundness of the mean free path concept.

Further support for this point of view is provided by the fact that the departures from Maxwell's prediction at very low pressures can also be given a reasonable interpretation. These departures arise because *the mean free path is becoming comparable with the dimensions of the apparatus,* in which circumstance collisions with the walls become the controlling factor rather than collisions between the molecules, and such concepts as viscosity and thermal conductivity cease to be applicable. A gas at such pressures is often known as a *Knudsen gas,* in recognition of **Knudsen's** extensive work on the behaviour of gases under "vacuum" conditions. The theory of these phenomena is usually quite easy if extreme "Knudsen" conditions are assumed, but it is much more difficult to deal with the intermediate range of pressures. We have already met the changes that occur in the formulae governing *thermal transpiration* as the pressure is altered, and similar changes occur in the formula governing the flow of gas in tubes. At high pressures, the gas obeys Poiseuille's law (p. 149), and the gas behaves just like a viscous liquid, the flow being proportional to the fourth power of the radius of the tube. We shall now investigate the flow on the assumption, due to **Knudsen**, that, during its travel across the tube, a typical molecule is accelerated by the pressure-gradient down the tube, but loses this drift velocity completely on collision with the wall. This would be the case if the surface of the wall were molecularly rough, so that the direction of travel of a molecule after reflection bears no relation to its direction on incidence. Actually, this assumption is an over-simplification, and a molecule colliding with the wall may retain a little of its drift velocity. The idea of a *slip coefficient* has been introduced by some workers to allow for such a possibility. This coefficient must not be confused with the *accommodation coefficient*, defined below (p. 134). On the assumption that no drift velocity is retained after a collision, we may calculate the rate of flow through a tube of radius R as follows: the distance a typical molecule travels between collisions with the walls is of the order of magnitude $2R$ (this being the figure for a molecule leaving a wall at right angles, but the value for other angles might be greater or less than this according to the direction of motion. In an accurate theory, we should have to average over all angles of inclination). The time between collisions is then $2R/c$, during which the molecule is accelerated by the pressure-gradient $\partial p/\partial y$ and acquires an extra velocity

$$= \frac{2R}{\bar{c}} \cdot \frac{1}{\rho} \frac{\partial p}{\partial y}$$

along the length of the tube, which is destroyed at the next collision with a wall. The average velocity along the tube is therefore half this, so that the mean drift velocity is $\frac{R}{\bar{c}} \cdot \frac{1}{\rho} \frac{\partial p}{\partial y}$. The mean volume of gas transported per second is then

$$= \frac{\pi R^3}{\rho \bar{c}} \quad \frac{p_1 - p_2}{L},$$

where $p_1 - p_2$ is the pressure-difference between the ends of the tube, and L is its length. This formula differs from Poiseuille's by a factor proportional to l/R, where l is the mean free path, and the flow is independent of the mean free path l, once this has become large compared with R. We might thus expect the following type of formula to be applicable to the intermediate case

$$Q = \frac{p_1 - p_2}{L} \quad \frac{R^3}{\rho \bar{c}} \left(A + \frac{BR}{l} \right) \ldots\ldots\ldots\ldots \quad (20)$$

where the constants A and B can be adjusted to agree with the Poiseuille type of flow at the high-pressure end, flow proportional to R^4/η, and to the Knudsen type at the low-pressure end, flow proportional to $R^3/\bar{c}$, the constant A of course depending on the slip coefficient. This is found to give a satisfactory interpolation formula. An analogous type of correction was made by **Millikan** in his experiments on the motion in air of very small oil-drops under an electrostatic field of force or gravity, for the determination of the electronic charge. He found that when a, the radius of a drop, became comparable with l, the drops tended to fall faster under gravity than was indicated by Stokes' law (p. 154) for the fall through a *continuous* viscous medium, the granular structure of matter becoming important. He found it possible to describe the departure from Stokes' law by means of a factor of the type $(1 + Ca/l)$ analogous to that occurring in equation (20).

Equation (20) is of great importance in the design of pumping apparatus for producing high vacua. It shows that, if full advantage is to be taken of the high pumping speed of powerful pumps, the tubes connecting them to the apparatus must be wide and short, and must contain no constrictions. If these precautions are neglected, it may easily happen that, for a given pumping speed Q, $p_1 - p_2$ in equation (20) becomes comparable with the pressure (perhaps a small fraction of a millimetre of mercury) that one is trying to reach, in which case the pumping speed is limited by the dimensions of the connecting tube rather than, as it should be, by the design of the pump.

In deriving equation (20) we appear to have neglected intermolecular collisions altogether at very low pressures. We are indeed assuming that a typical molecule collides with the walls far more often than with other molecules, but this is not at all the same thing as neglecting the latter type of collision altogether. When we speak of molecules being "accelerated along the tube by the pressure-gradient" between collisions with the walls, we are really using an idea that is more appropriate to a continuous medium, and it is not easy to see how a typical molecule flying *across* the tube can be accelerated *along* it, unless it makes at least one collision with another molecule. This is a typical illustration of how "hydrodynamic" concepts, appropriate to a continuous medium, break down for a Knudsen gas. The situation could be rather better stated in this way: a gradient of *pressure* down the tube implies a corresponding gradient of *density* (whether the mean free path is long or short), and the effect of intermolecular collisions is to secure that there is a net flow of gas from the higher density to the lower density regions, at exactly the rate corresponding to the pressure-difference. Put in another way, the random velocities of the molecules and their mutual collisions ensure that there will be a net *flow of gas* from regions of higher *density* to regions of lower *density*. At the same time, the collisions with the *walls* give rise to the *pressure*. In an ordinary gas or liquid, we are accustomed to think of the pressure-difference as being the "cause" of the flow from one place to another, but it is equally possible to think of the flow as being the "cause" of the pressure-difference, and this point of view is definitely preferable for the Knudsen gas. The pressure-difference between the ends of the tube is fixed by the principle of conservation of momentum, once the rate of flow has settled down to a steady value.

A very similar problem is to determine the force acting on a surface moving tangentially. If the velocity of the surface is U along the x-axis, then we are to suppose that the molecules striking it have, before impact, an average velocity component along the x-axis of zero, whereas, after striking it, they have an average velocity component of U in this direction. The number of molecules falling per second on unit area of surface is $\frac{1}{6}N\bar{c}$, and the tangential force on unit area is, by conservation of momentum, $\frac{1}{6}Nm\bar{c}U$.

This result can be applied to the *quartz fibre manometer*. This consists simply of a stretched quartz fibre, and a pressure measurement is taken by setting it into vibration and determining the logarithmic decrement of its oscillations. Since the force is proportional to the velocity, the theory of the damped harmonic oscillator (p. 285) is applicable, according to which the logarithmic decrement is proportional to the damping force. Under "Knudsen" conditions, the mean free path of the molecules is large compared with the radius of the fibre and the probability of the same molecule colliding a second time with the fibre before making many collisions with the walls is very small, so that none of the momentum imparted to the gas ever returns to the fibre. (Under ordinary conditions of short mean free path some of it would.)

In such circumstances we may assert that

$$\text{log-decrement} = A + BNm\bar{c}$$

or, since

$$m\bar{c}^2 = \tfrac{3}{2}kT$$

$$\text{log-decrement} = A + NB'\sqrt{mT} \quad \ldots\ldots\ldots\ldots \quad (21)$$

where A represents the damping due to friction in the quartz fibre itself. The constant B' is best determined by calibration of the manometer against a MacLeod gauge. It must, however, be remembered that, in a manometer of this kind, the presence of mercury or other *vapour* in the apparatus will also result in a contribution to the damping of the fibre, but such a pressure will not contribute to the recording of the MacLeod gauge.

A second application is the Dushman molecular gauge. This consists of two parallel discs, one of which can be rotated by a special type of synchronous motor, while the other is suspended on a torsion fibre. Under "Knudsen" conditions, assuming a slip coefficient of zero, one may safely assume that nearly all the momentum carried away from the rotating disc by the mechanism outlined above will be transferred to the other disc, provided that the discs are spaced at a distance small compared with the mean free path. The constants of this gauge can, in principle, be predicted theoretically if edge effects are eliminated by means of the well-known "guard-ring" principle, the suspended disc being surrounded by a fixed ring in its own plane (Fig. 50).

The radiometer. The accommodation coefficient

It is now natural to ask how phenomena involving heat-transfer are affected at very low pressures. The fundamental mechanism becomes extremely simple: we are to think of a typical molecule striking a heated surface and rebounding with an increase of its kinetic energy, some of which increase may be transferred to the next surface it strikes. However, it does not necessarily follow that an average molecule rebounding from a hot surface has a kinetic energy corresponding precisely to the temperature of that surface. Knudsen introduced the conception of the *accommodation coefficient* defined in this way: let a typical molecule from a gas at a temperature T strike a surface at a temperature T_1. Let T' be the temperature corresponding to the average energy of such a molecule on leaving the surface. We then define the *accommodation coefficient* as the ratio $(T' - T)/(T_1 - T)$, which may vary between zero and unity. A typical value is 0·5 for a metal surface covered with adsorbed gas, but **Roberts** showed that the value for a *clean* metal surface may be very much smaller (of the order of 0·05). This quantity is quite distinct from the *slip coefficient* (p. 131), which is concerned with *velocities* rather than *temperatures*, and is often assumed to be zero, though it is difficult to measure accurately. The accommodation coefficient can, however, be measured directly, the best method being to stretch a fine wire along a cylindrical tube, heat the wire electrically, simultaneously measuring its resistance by means of a Wheatstone-bridge circuit. One can thus get a relation between the temperature and the rate of loss of heat from the wire. If the pressure is low enough for the mean free path to become long compared with the diameter of the wire, a typical molecule leaving the wire will make a great many collisions with the wall, and will therefore possess a

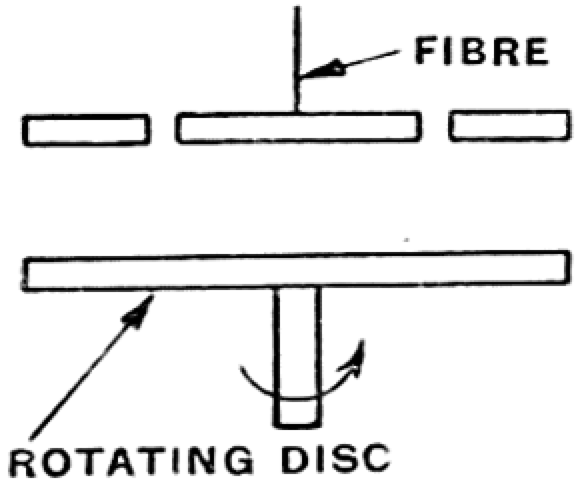


FIG. 50.

mean kinetic energy corresponding to that temperature, before it strikes the wire again. Since we know the number of molecules striking unit surface of the wire at a given temperature and density, we have a direct measure of the energy lost by the wire to an average molecule, and so of the accommodation coefficient.

If a molecule acquires extra energy on rebounding from a heated surface, and we have "Knudsen" conditions, so that the molecule strikes the walls several times before reaching the surface again, it follows, by conservation of momentum, that there must be a net force normal to the surface. For if c_1 is the mean velocity corresponding to the temperature of the rest of the apparatus, while c_2 is the mean velocity of the molecules on leaving the surface, the momentum imparted to the surface per unit area per unit time is proportional to $mNc_1(c_2 - c_1)$. If T is the temperature of the apparatus, corresponding to c_1 and $T + \Delta T$ the temperature of the surface, then, by the definition of the accommodation coefficient a, the temperature corresponding to c_2 is $T + a\Delta T$. For small ΔT, the force on the surface is therefore proportional to

$$mNT^{\frac{1}{2}}[(T + a\Delta T)^{\frac{1}{2}} - T^{\frac{1}{2}}] \simeq \tfrac{1}{2}mNa\Delta T.$$

This effect is the basis of the *Crooke's radiometer*, which makes evident the net force on a vane, one side of which is polished, the other blackened. If this is exposed to radiation the blackened side becomes hotter than the polished side and there is a larger force on the blackened side. By mounting several such vanes on a "spider" freely pivoted, we get an instrument that can be used to detect thermal radiation, though it is not often used as a *measuring* instrument. The effect becomes small at very low pressures because of the small value of N, and practically disappears at high pressures because the mutual collisions between the molecules secure a practically uniform temperature in the surrounding gas on both sides of the vane. The effect is thus a typical "Knudsen" one, requiring a long mean free path before it becomes manifest. It must not be confused with *radiation pressure*, which (i) is a physical effect of the radiation itself, (ii) is of a smaller order of magnitude, (iii) is nearly twice as great on a *polished* surface as on a *blackened* one, and (iv) exists quite independently of whether any gas is present or not.

The radiometer effect has been exploited by **Knudsen** in his so-called *absolute manometer*. Its principle is illustrated in Fig. 51. A, B are two plates electrically heated, placed on opposite sides of a "picture-frame" C suspended on a torsion fibre, so that the repulsion that A or B exerts on C due to the increase in mean velocity of the molecules leaving A and B can be inferred. The name absolute manometer arises from the fact that the force is *independent of the nature of the gas* and therefore measures the total pressure of all gases and vapours present. It does, however, require knowledge of the accommodation coefficient of the surfaces, but this can be determined by a separate experiment of the type already described (p. 135), or the gauge can be treated as non-absolute and calibrated against a MacLeod gauge, due account being taken of the possible pressure of vapours, as mentioned above (p. 133).

6. Concluding remarks

Enough has been said to prove the very great value of the concept of the mean free path. Like many other useful ideas, it is clear enough intuitively, but it becomes extremely difficult to write down a precise definition that shall cover every possible case. There is no particular difficulty about extending the mean free path concept to cover cases, such as gases at very high or very low temperatures, in which a quantum-mechanical treatment of molecular collisions becomes advisable. The results of such calculations are nearly always presented as a *collision cross-section.* Having defined a particular type of collision, resulting in, say, the deflection of the path of the colliding atom through more than a certain angle, it is then possible to define a "target area" associated with each molecule, which must be "hit" by another molecule in order that a collision of the specified type may take place. We have already seen how closely the classical notions of "target area" and mean free path are related (p. 123).

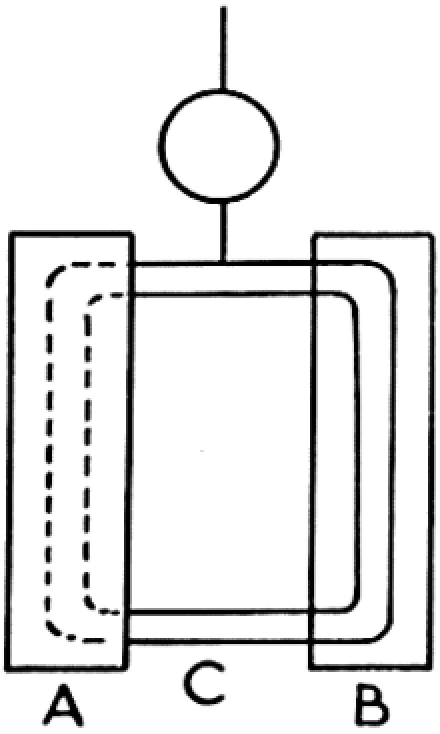


FIG. 51.

One of the main reasons why the kinetic approach to liquids has not led to any very spectacular advances in theory is that the mean free path concept ceases to have any real utility. In the first place the "free volume" in liquids and solids is always comparable with the actual volume of the molecules themselves. This means that collisions occur each time a molecule travels more than about a diameter, and the "mean free path" hardly changes with pressure and temperature. Secondly, a simultaneous collision of three or more molecules is a comparatively rare event in a gas, and such collisions can safely be neglected, but this is no longer so in the much denser liquids and solids, in which simultaneous collisions of three or four molecules are nearly as frequent as binary ones.

BIBLIOGRAPHY FOR CHAPTER V

J. K. Roberts, *Heat and Thermodynamics*. 4th Edn. Blackie, London and Glasgow (1951).

J. H. Jeans, *Dynamical Theory of Gases*. C.U.P. (1921).

J. H. Jeans, *Introduction to Kinetic Theory of Gases*. C.U.P. (1940).

L. B. Loeb, *Kinetic Theory of Gases*. 2nd Edn. McGraw-Hill, New York (1934).

F. H. Newman, *Production and Measurement of Low Pressures*. Benn, London (1928).

M. Knudsen, *Kinetic Theory of Gases, Methuen Monograph*. 3rd Edn. London (1950).

J. C. Maxwell, *Scientific Papers*, Vol. I. C.U.P. (1890).

J. K. Roberts, *Some Problems in Adsorption*. Cambridge Physical Tracts (1939). Revised by A. R. Miller (1951).

Examples on Chapter V

1. Calculate the root-mean square velocity of hydrogen molecules at N.T.P., outlining any assumptions you make and the derivation of any formulae you need. (Mass of hydrogen molecule, 3×10^{-24} g.; Avogadro's number, 6×10^{23} mol./g.-mol.; gas constant, $8{\cdot}5 \times 10^7$ ergs/mol./degree C.)

2. Distinguish carefully between the concepts of "mean free path" and "mean molecular distance" in a gas. Derive a rough formula for the ratio of these two quantities in a gas. Is it ever possible for this ratio to approach unity?

3. Explain how one can derive a theoretical value for the viscosity of a gas assumed to consist of rigid spheres of known diameter σ. Outline the derivation of any formulae you use.

4. "The concept of the viscosity of a gas ceases to have a definite meaning at very low pressures." Discuss this statement, qualifying it in any way you think necessary.

5. A pump of fixed pumping speed is to be used to evacuate apparatus to a pressure of 10^{-5} mm. of mercury. What considerations determine the design of the tubes used to connect the apparatus to the pump?

6. Derive from first principles a relation between the viscosity and thermal conductivity of a gas.

7. Describe and explain the phenomenon of thermal transpiration at low pressures, illustrating your answer by reference to: (*a*) vapour pressure measurements on a liquid such as hydrogen at a temperature well below its boiling point; (*b*) the measurement of the pressure in an evacuated vessel. In what way is the phenomenon modified as the pressure rises?

CHAPTER VI

HYDRODYNAMICS

1. Introduction

In Chapter IV we have established a number of results about the behaviour of solids, on the assumption that they are continuous media, described by quantities such as density and the elastic moduli, which are treated as varying continuously from point to point. We have indicated roughly how such conceptions may be justified on the basis of the kinetic theory of matter, and, in Chapter V, we have shown how the properties of gases, not too near the liquid region, can also be interpreted satisfactorily on a kinetic basis. In the present chapter we shall see how far the properties of a liquid can be interpreted on a "continuum" basis, and, in Chapter VII, we shall examine some of the attempts that have been made to build up a kinetic theory of liquids corresponding to the kinetic theory of gases.

2. Definition of a liquid. Viscosity

As usual, it turns out to be a difficult matter to set up a logically precise definition of a liquid. Perhaps the best and most usual definition is that a liquid is incapable of permanently resisting a shearing stress, though it does resist compressive stresses, and, as we shall see in Chapter VII, tensile stresses also. This means that it does not, like a gas, at once *fill* any vessel in which it is placed, though it will, in the course of time, accommodate itself to the *shape* of the vessel. The time taken by a liquid to settle down to its final shape, after transfer to another vessel, is a very important property. This time may vary from a fraction of a second for mobile liquids such as water and ether, through minutes for such liquids as treacle, glycerine, or heavy oil, to hours and days for such bodies as pitch and cobbler's wax, which are, for most practical purposes, solids, though they must be regarded as liquids according to our strict definition.

Shear viscosity.—It is customary to describe the yielding of a liquid to a shearing stress by means of the *coefficient of shear viscosity*, which is defined as the *tangential force per unit area acting between two layers of fluid per unit velocity gradient in a direction perpendicular to the layers.* This definition is closely analogous to the definition of the rigidity modulus of elasticity (p. 81) the difference being simply that, in the definition of viscosity, *rate of change of shear* replaces the shear itself which occurs in the definition of rigidity modulus. This definition is equally applicable to a liquid or gas.

Expansion viscosity.—The question of what happens when the rate of change of shear associated with the tangential motion is nearly zero now arises. An example of this situation is a sound-wave so far from its source that it may be treated as nearly plane. There is, in such a wave, a difference in stream velocity along the direction of propagation of the wave, that is, in the same direction as the stream velocity itself. If we refer to the derivation of shear viscosity in the kinetic theory of gases (p. 124) we find that viscosity is to be attributed to the transport of drift momentum by the molecules. This transport of momentum takes place in all directions, in particular such transport occurs *along* the direction of propagation of a sound-wave, from the regions of high to regions of low stream velocity, and thus tends to damp out the sound-wave. In this way we are led to the idea of a coefficient of *expansion viscosity*, defined as *the normal force per unit area per unit velocity gradient acting between two layers of fluid in a direction perpendicular to the layers.* It may also be expressed, by analogy with the *bulk modulus* of elasticity, as

$$\frac{\text{Pressure opposing expansion or contraction}}{\text{Fractional rate of change of density per unit time}}.$$

Maxwell proved the following relationship between the two coefficients of viscosity in a gas on the basis of kinetic theory.

$$\frac{\text{Expansion viscosity}}{\text{Shear viscosity}}=\frac{1}{3}$$

As we shall see in the next chapter, the mechanism causing viscosity of a liquid is entirely different from that acting in a gas, and there is no reason to believe in such a relationship as the above, the available evidence definitely points in the other direction. In this book, unless another meaning is obvious, the term "viscosity" is to be taken to mean shear viscosity, and it is assumed that the relation is obeyed unless the contrary is stated. It should be mentioned that some writers speak of a fluid obeying this relation as having "zero coefficient of second viscosity", but it has seemed to the writer more natural to define two coefficients of viscosity analogous to the two coefficients of elasticity. There does not seem to be any "viscous analogue" of Poisson's ratio.

The concept of viscosity cannot be interpreted physically except in terms of some molecular hypothesis. Nevertheless, one cannot develop hydrodynamics very far without introducing such a concept. The consequences of assuming that the viscosity of a fluid vanishes can, in many cases, be worked out exactly, but they are, in several respects, utterly unrealistic. (We conclude, for example, that no force opposes a solid body moving steadily through a liquid.) On the other hand, there are quite a number of results in hydrodynamics which are independent of the numerical magnitude of the viscosity, and hold, provided only that the viscosity is "small" in a sense that will be explained more fully later (p. 154), but these results cannot be deduced if one assumes that the viscosity is rigorously zero to begin with.

Non-Newtonian fluids.—From the kinetic theory of *gases* we can deduce that the stresses should be proportional to the rates of deformation, at all events up to velocities comparable with the (very high) velocities of thermal agitation. A fluid showing such proportionality is usually described as *Newtonian*, a fluid that does not is *non-Newtonian*. Since the mechanisms producing viscosity in a liquid are still rather uncertain, one cannot say *a priori* whether a *liquid* should be Newtonian or not. In fact, nearly all pure liquids are Newtonian, but many colloids, solid-liquid mixtures, and solutions of long-chain molecules, are non-Newtonian. The lack of linearity may be in either direction, the stress may increase either more, or less, rapidly than the velocity-gradient. The latter behaviour occurs in quicksand (roughly a sand-clay-water mixture), and is known as *thixotropy*.

It is seldom necessary to consider anisotropy in the mechanical properties of a liquid, though it can occur with plate-like or with very long and thin molecules, an example being a solution of tobacco mosaic virus in water. Such systems are sometimes called "liquid crystals", but we shall not deal with them in this book.

3. The equations of hydrodynamics

In hydrodynamics we use the words "particle" and "particle-velocity" in a special sense. They refer to volumes of fluid whose dimensions are physically small compared with other lengths entering into the theory, but which are, nevertheless, large enough to contain a great many molecules. The "particle velocity" is *not* the velocity of a single molecule but the average over a great many molecules, that is to say it corresponds to the *drift velocity* introduced into the kinetic theory of gases (p. 122). A term that is also often used to describe this velocity, the velocity of translation of a small volume of fluid, is *stream velocity*, which is treated as varying continuously from place to place.

We arrive at the equations of hydrodynamics, first by writing down the condition that fluid is conserved, any change in the mass of fluid within a small volume being equated to the total flux into or out of that volume, secondly, by equating the mass times acceleration of the small volume to the total force acting on it. These forces are of three distinct kinds:

(*a*) They may arise from the direct action of an external field of force on the particles. The most usual force of this kind is that of gravity, but other external forces, *e.g.* electrostatic, are equally effective.

(*b*) They may arise as a result of hydrostatic pressure, which is simply another way of describing the mutual forces between neighbouring elements of liquid, supposed to be at rest relative to one another. One must distinguish very carefully between the gravitational force acting on a particle of fluid because of its weight, and the forces exerted on the particle by the particles surrounding it. For hydrostatic equilibrium, the net force on any particle is zero, the effects of gravitation and hydrostatic pressure balancing, but this need no longer be so if the particle is in motion.

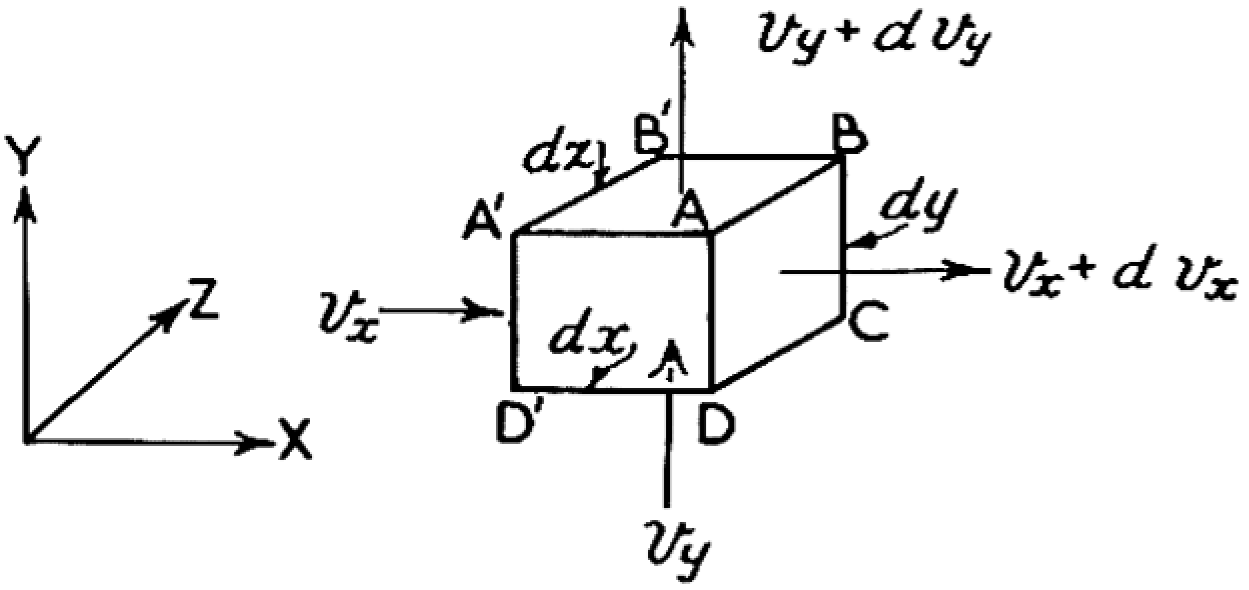


FIG. 52.

(*c*) They may arise as a result of viscous forces between layers of fluid in relative motion.

The equation of continuity.—Let Fig. 52 represent a small cuboid with sides dx, dy, dz. The rate of increase of the mass contained within the volume is clearly $\frac{\partial\rho}{\partial t}.dx\,dy\,dz$, where ρ is the mean density. The flux across the surface $A'B'C'D'$ *into* the volume is $\rho v_x\,dy\,dz$, where v_x is the stream velocity in the x-direction averaged over this small area. The flux across $ABCD$ *out of* the volume is $\rho v_x\,dy\,dz$, and the difference in the flux across the two surfaces must either remain in the volume, or else pass across one of the other four surfaces. This difference arises because ρ and v_x may both vary from point to point. (This is to be taken in the macroscopic sense, and does not mean that the molecular structure of the fluid is being considered.) The difference between ρv_x at the planes $A'B'C'D'$ and $ABCD$ is $-\frac{\partial}{\partial x}(\rho v_x)\,dx$, by the definition of a differential coefficient. There will be similar contributions to the net rate of flow of matter into the volume from the other two pairs of faces, so that we have finally

$$\frac{\partial \rho}{\partial t} dx\,dy\,dz = dx\,dy\,dz\left[-\frac{\partial}{\partial x}(\rho v_x) + \text{two similar terms}\right]$$

or

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x}(\rho v_x) + \frac{\partial}{\partial y}(\rho v_y) + \frac{\partial}{\partial z}(\rho v_z) = 0 \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (1)$$

We must be careful to notice the exact meaning of the velocity components v_x, v_y, v_z. In the customary way of formulating hydrodynamics, the *Eulerian formulation*, these quantities refer to the components of the stream velocity *at a given point in space.* The motion of the fluid is completely described by the variation of these quantities with time at each point in space. This variation is *not* the same thing as the history of a particular "particle" of fluid, because, in general, a given particle moves about in the fluid, whereas v_x, etc., refer to the local stream velocities *at a fixed position in space.* This difference can often be neglected, but it is necessary to attend to it when we are writing down the acceleration equation.

It is possible to formulate hydrodynamics in such a way that the history of all the individual "particles" is traced as they move about in the fluid. This is called the *Lagrangian formulation* of hydrodynamics. It is mathematically equivalent to the Eulerian formulation, but the resulting equations are usually cumbersome. These are, however, certain special problems for which Lagrange's method is simpler, but we shall not meet any of them in this book.

The acceleration equation.—We equate the net force on a particle of fluid to its mass times acceleration, and, in writing down the acceleration we encounter a difficulty not unlike that which occurs when we attempt to write down the radial and transverse accelerations of a particle moving in an orbit (p. 21). In the former case the difficulties were caused by the fact that what we wanted was the change in velocity *along a fixed direction.* Here, what we want is the change in velocity *associated with a particular particle* of fluid. We may safely equate the velocity of a particular particle to the local velocity at that point in the fluid, but we may not say that the acceleration of a particular particle is $\partial v_x/\partial t$. For in a small time dt, our particle has travelled a distance dx along the X-axis, dy along the Y-axis, etc., and our change in velocity of that particle in the time dt is given by $v_x(x + dx, y + dy, z + dz, t + dt) - v_x(x, y, z, t)$, where dx, dy, etc., are not independent of dt, but are so chosen that x, $x + dx$, are successive X-coordinates of the same particle. The condition for this is $dx/dt = v_x$, and similar relations hold for dy and dz.

Thus, the acceleration is to be calculated from the following expression

$$a_x = \underset{dt \to 0}{Lt} \frac{1}{dt}\left[\frac{\partial v_x}{\partial t} dt + \frac{\partial v_x}{\partial x} dx + \frac{\partial v_x}{\partial y} dy + \frac{\partial v_x}{\partial z} dz\right]$$

where the expression in the square brackets represents the total change in velocity due to varying all four independent variables. This expression is quite general, and is a direct deduction from the definition of differentiation. If now we want the difference in the successive velocities *of a particular particle*, the ratio of dx to dt must be properly chosen, must in fact be equal to v_x, and so on.

Thus we arrive at the following expression for the acceleration

$$a_x = \frac{\partial v_x}{\partial t} + v_x \frac{\partial v_x}{\partial x} + v_y \frac{\partial v_x}{\partial y} + v_z \frac{\partial v_x}{\partial z}.$$

Applying Newton's second law to a small volume of sides $dx\, dy\, dz$ we have

$$\rho\left(\frac{\partial v_x}{\partial t} + v_x \frac{\partial v_x}{\partial x} + v_y \frac{\partial v_x}{\partial y} + v_z \frac{\partial v_x}{\partial z}\right) dx\, dy\, dz =$$

$$\left(F_x - \frac{\partial P}{\partial x} + \text{viscous force}\right) dx\, dy\, dz \quad \ldots\ldots\ldots\ldots \quad (2)$$

where F_x is the external force acting in the X-direction on unit volume. The term $\partial P/\partial x$ arises because, if there is a change in the pressure as x varies, there will be a force $P\, dy\, dz$ acting on the area $A'B'C'D'$ (Fig. 52) of the small element, while there will be a force $\left(P + dx \frac{\partial P}{\partial x}\right) dy\, dz$ acting backwards on the area $ABCD$. The general derivation of the viscous force is complicated and the reader is referred to Lamb's *Hydrodynamics* for details. It is better to derive *ab initio* the expression applicable to each particular problem, which can usually be done quite easily directly from the definition of viscosity.

Equations (1) and (2) are not theoretically sufficient to determine the motion of a system. It is necessary to assume the equation of state, some relation between ρ, the density, and P, the pressure. Two assumptions frequently used are:

(*a*) *The "incompressible" assumption.*—ρ is supposed to be independent of P, a supposition that is nearly true for many liquids.

(*b*) *The "acoustic" assumption.*—ρ is supposed to be a linear function of P. All substances have a range of values of P for which this is true.

4. Some solutions of the equations of hydrodynamics

We now have three equations of type (2) together with equation (1), and we have four dependent variables (v_x, v_y, v_z, and p or ρ), but even now the solution is not determinate. As is well known, the unique solution of a set of differential equations requires a careful specification of the initial and boundary conditions. When we do this, it often turns out that an actual problem is too complicated to solve. The tendency, therefore, has been rather to compile "collections" of particular solutions of the equations, and to use them as a means of gaining physical insight into more complicated problems. In this respect hydrodynamics resembles stress analysis, electromagnetics, and other branches of "field-physics." We shall now examine a number of such simple solutions.

A. Hydrostatics

This case is the simplest of all, the particles being at rest. We then have $\partial P/\partial x = F_x$, $\partial P/\partial y = F_y$, $\partial P/\partial z = F_z$. If the external field is gravity, acting in the negative z-direction, we have $F_x = F_y = 0$, $F_z = -\rho g$. We thus obtain the well-known law of hydrostatics that *pressure depends only on the depth,* which we have arrived at by considering the equilibrium of a particle of fluid under its weight and the resultant hydrostatic force. For a liquid, ρ is practically constant and $\partial P/\partial z = -\rho g$ implies that $P = P_0 - \rho g z$, that is *pressure increases linearly with depth.* For a gas, the relation between density and pressure is $P = NkT\rho$, where N is the number of molecules per unit mass, so that $NkT\dfrac{\partial \rho}{\partial z} = -\rho g$, which integrates into

$$\rho = \rho_0 e^{-gz/NkT} = \rho_0 e^{-mgz/kT},$$

where m is the mass of a molecule. This relation has already been derived by other means for a gas under gravity (p. 119).

B. Flow in one dimension

Let us assume that v_y and v_z are zero. Equations (1) and (2) now read:—

$$\frac{\partial(\rho v_x)}{\partial x} + \frac{\partial \rho}{\partial t} = 0$$

$$\frac{\partial v_x}{\partial t} + v_x \frac{\partial v_x}{\partial x} + \frac{1}{\rho}\frac{\partial P}{\partial x} = \frac{F_x}{\rho}.$$

For a liquid, ρ is practically constant. It will be noticed that the rate of increase with time of, *e.g.*, the velocity of flow of a liquid along a pipe is not, as we might at first expect, proportional to $\frac{F_x}{\rho} - \frac{1}{\rho}\frac{\partial P}{\partial x}$ but to

$$\frac{F_x}{\rho} - \frac{1}{\rho}\frac{\partial P}{\partial x} - \frac{\partial}{\partial x}\left(\frac{v_x^2}{2}\right).$$

Thus, even in a liquid with negligibly small viscosity, a given pressure gradient apparently becomes less effective in accelerating a liquid as the velocity of the latter increases. This is one aspect of a very general phenomenon known as the *Bernoulli effect*, after its discoverer. We shall meet other instances of this effect, which arises from the fact that the acceleration of a particular particle of fluid is not precisely equal to $\partial v_x/\partial t$.

Bernoulli's theorem

If we restrict the motion of our liquid to be *irrotational*, that is, if we impose on it the conditions, which we shall discuss later,

$$\frac{\partial v_z}{\partial y} - \frac{\partial v_y}{\partial z} = \frac{\partial v_x}{\partial z} - \frac{\partial v_z}{\partial x} = \frac{\partial v_y}{\partial x} - \frac{\partial v_x}{\partial y} = 0$$

or curl $\bar{v} = 0$,
then equation (2) may be written

$$\frac{\partial v_x}{\partial t} + v_x\frac{\partial v_x}{\partial x} + v_y\frac{\partial v_y}{\partial x} + v_z\frac{\partial v_z}{\partial x} + \frac{1}{\rho}\frac{\partial P}{\partial x} = \frac{F_x}{\rho} \quad \ldots\ldots\ldots (3)$$

with two similar equations. For steady flow, without external forces other than gravity

$$\frac{\partial v_x}{\partial t} = 0 \text{ and } F_x = 0,\ F_y = 0,\ F_z = -\rho g.$$

Equation (3) can then be integrated if ρ is constant and the result is

$$P + \tfrac{1}{2}\rho\,(v_x^2 + v_y^2 + v_z^2) = \text{constant} - \rho g z \quad \ldots\ldots\ldots (4)$$

This means that, in a liquid not subject to external forces, the pressure is least where the velocity is greatest. If, for example, a liquid is flowing, *e.g.* through a pipe with a constriction, the velocity has to be greatest at the constriction, and the pressure is accordingly least there. This result is true also for compressible fluids and is sometimes known as *Hawksbee's law*. It is the basis of the vacuum cleaner, the filter pump, and the scent-spray. It is also exploited in the *Pitot tube*, used for measuring stream-velocity. This consists essentially of a fine pointed tube, directed as nearly as possible along the stream-lines. In equilibrium, no fluid enters or leaves it, and the pressure inside it exceeds that in the neighbouring liquid or gas by just the term $\frac{1}{2}\rho v^2$. If the compressibility of the fluid is to be taken into account, equation (4) is replaced by

$$\int dP/\rho + \tfrac{1}{2}(v_x^2 + v_y^2 + v_z^2) = \text{constant} - gz \; \ldots\ldots \quad [4(a)]$$

A particular case of Bernoulli's theorem is *Torricelli's theorem.* Suppose we have a liquid escaping from a vessel through a small aperture, so that the flatness of the free surface is not disturbed. In equation (4), the velocities near the free surface are very small and the constant is P_0, the pressure at the free surface, if we take $z = 0$ to be this surface. At the orifice at depth z, $P = P_0$ also, so that equation (4) gives us $\frac{1}{2}v^2 = -gz$, i.e. *the velocity of emergence is just that corresponding to free fall through a distance equal to the depth of the orifice below the free surface.*

C. Liquid in a rotating vessel

We determine the form of the free surface in a vessel rotating with a constant angular velocity ω about the z-axis. It makes no difference to the results whether we impose the condition that the accelerations $\partial v_x/\partial t$, $\partial v_y/\partial t$ must be those appropriate to motion in a circle of radius $(x^2 + y^2)^{\frac{1}{2}}$, or whether we change over to axes rotating with the vessel, and introduce "centrifugal force terms", to allow for the effect of this rotation. We consider the variation of the pressure as a function of x and z. Because of the rotational symmetry, it is sufficient to consider the case $y = 0$. A particle at a distance x from the axis (Fig. 53) will have, under steady conditions, a velocity $x\omega$, along the y-axis. The acceleration towards the axis is $x\omega^2$, along the x-axis. Substituting in equation (3) and putting $F_x = F_y = 0$, $F_z = -\rho g$ we get $-\omega^2 x + \dfrac{1}{\rho}\dfrac{\partial P}{\partial x} = 0$,

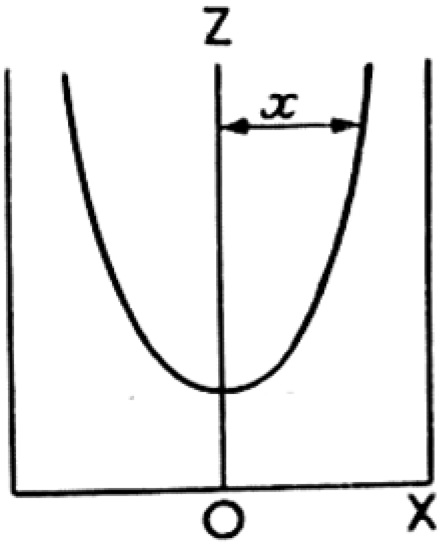


FIG. 53.

$$\frac{\partial P}{\partial z} = -\rho g,$$

which integrate to give us $P + \rho g z - \rho\dfrac{\omega^2 x^2}{2} = \text{constant}.$

The pressure is therefore constant along the curves

$$gz = \omega^2 x^2/2 + \text{constant}.$$

Since the pressure is necessarily constant along a free surface, the free surface must be of this parabolic form. Because of the symmetry between x and y, the form of the free surface will be that of a parabola rotated about its axis. This form is independent of ρ, the density.

D. Flow of a viscous liquid in a tube. Poiseuille's law

If a viscous liquid flows through a uniform tube under a not too large pressure-head, it can be shown, *e.g.* by introducing small quantities of dye into parts of the liquid, that the flow is along the length of the tube. If the pressure-head is increased, this ceases to be true, the lines of flow first becoming sinuous and finally disappearing altogether, a typical fluid "particle" dashing from side to side of the tube in an apparently random fashion. This latter state of affairs is known as *turbulent flow*, which will be discussed later (p. 152). It has many points of analogy with molecular motion, but is very much more difficult to study mathematically, because, in the kinetic theory of gases, it is permissible to think of a *small* drift velocity superimposed on a random distribution of velocities of agitation, whereas, in turbulent flow, the drift velocity and random velocity are almost always comparable in magnitude.

The low velocity regime is capable of exact mathematical treatment, and is called the region of *laminar flow*. If we assume, as is justified by the experiments, that the flow is entirely along the length of the tube, we may, once conditions have become steady, equate the viscous force acting on a small element of liquid to the force associated with the gradient of pressure. The element of liquid that we shall take is a hollow cylinder (concentric with the tube), of length dx, inner radius r, and outer radius $r + dr$. We also assume that the fluid is stationary at the walls, which is probably always correct for a liquid, but is less certain for a gas (see p. 131). The viscous force acting on the outer surface of the element is, by the definition of viscosity, equal to the viscosity × area of surface × velocity gradient, that is

$$\eta \times 2\pi (r + dr)\, dx \times \left(\frac{\partial v}{\partial r}\right).$$

Since v is decreasing as r increases, this force is tending to hold the element back. The viscous force acting on the inner surface is, on the contrary, tending to drag the element forward in the direction of v, because the fluid is moving faster near the centre of the tube. The net decelerating force has, therefore, to be equated to the force on the element due to the pressure in the tube, which is $-\left(\frac{\partial P}{\partial x}\right) dx \times$ area of end of hollow cylinder, that is $-\partial P/\partial x . dx . 2\pi r dr$. Equating these forces, we get

$$\left(\frac{\partial P}{\partial x}\right) . 2\pi r\, dr\, dx = 2\pi\eta\, dx \left[\left\{r\frac{\partial v}{\partial r}\right\}_{r+dr} - \left\{r\frac{\partial v}{\partial r}\right\}_{r}\right] \ldots\ldots \quad (5)$$

because, in calculating the net viscous force on the element, we have had to allow for the differing areas of the inner and outer curved surfaces of the element, as well as for a possible variation of $\partial v/\partial r$ with r. Equation (5) may also be written

$$\frac{\partial P}{\partial x} = \frac{\eta}{r}\frac{\partial}{\partial r}\left(r\frac{\partial v}{\partial r}\right) \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad [5(a)]$$

If we now suppose that our tube connects two large vessels full of fluid, the pressure must be constant over the cross-sections at both ends (unless the velocities of entry and exit are very large) and we must therefore suppose that $\partial P/\partial x$ is constant over every cross-section between the ends. Equation [5(a)] then shows that $\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial v}{\partial r}\right)$ must be independent of r. Let it be equal to A. Then, integrating we have

$$r\frac{\partial v}{\partial r} = \frac{Ar^2}{2}$$

(the constant of integration being zero, because any other value would imply an infinite value of $\partial v/\partial r$ at $r = 0$). Integrating again

$$v = \frac{Ar^2}{4} + B,$$

where B is a second constant of integration that must be chosen so that $v = 0$ at $r = a$, the wall of the tube. Thus, if v_0 is the velocity at the centre of the tube, the velocity at other radii must be given by the relationship

$$v = \frac{v_0}{a^2}(a^2 - r^2) \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (6)$$

Substituting this in equation [5(a)] we find

$$-\frac{\partial P}{\partial x} = \frac{4\eta v_0}{a^2} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (7)$$

We cannot, however, usually measure v_0 directly, but it is possible to measure F, the total *volume* of fluid flowing through the tube in unit time. The total rate of flow through a cross-section lying between circles of radii r and $r + dr$ is $v.2\pi r dr$. Substituting the value of v given by equation (6) and integrating between 0 and a, to cover the whole cross-section

$$F = \int_0^a v.2\pi r dr = 2\pi\int_0^a \frac{v_0 r\,(a^2 - r^2)}{a^2}\,dr$$

$$= \frac{\pi}{2}v_0 a^2 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (8)$$

Comparing (7) and (8) and eliminating v_0, we have finally

$$F = -\frac{\pi a^4}{8\eta}\frac{\partial P}{\partial x} \quad \text{..................} \quad (9)$$

that is, the total rate of discharge is proportional to a^4, and, as we might expect, inversely proportional to the viscosity. Equation (9) is known as Poiseuille's law. For an incompressible fluid, the total rate of *mass* flow is simply equal to ρF and the pressure gradient is constant, but for a gas it is necessary to take the compressibility into account. If P_1 and ρ_1 are the pressure and density in one vessel, we have by Boyle's law $P/\rho = P_1/\rho_1$. From the equation of continuity, it follows that ρF, the rate of mass flow, is constant along the length of the tube, and equal to Q, say. Then, from equation (9),

$$Q = \rho F = -\frac{\pi a^4}{8\eta}\frac{\rho_1}{P_1} P \frac{\partial P}{\partial x},$$

so that, for a gas, the pressure-gradient is not constant, but $P\dfrac{\partial P}{\partial x}$ or $\dfrac{1}{2}\dfrac{\partial (P^2)}{\partial x}$ is, so that we have

$$Q = \frac{\pi a^4}{16\eta}\frac{\rho_1}{P_1}\frac{(P_1^2 - P_2^2)}{l} \quad \text{................} \quad (10)$$

where P_1, P_2 are the pressures in the two vessels.

Many methods of measuring viscosity of gases rely on allowing the pressure in a large vessel to drop gradually as the gas flows out through a capillary tube and measuring the *time-interval* associated with the drop of the internal pressure P_1 from P_3 to a smaller value P_4, the external pressure meanwhile staying at a constant value P_2. If V is the volume of the vessel and Q is the mass of gas flowing out per second, then the rate of change of pressure is given by $\dfrac{1}{P_1}\dfrac{dP_1}{dt} = -\dfrac{Q}{\rho_1 V}$, because, for constant volume, the pressure is proportional to the total mass present, so the rate of fractional fall of pressure is equal to the rate of fractional loss of mass. Substituting in equation (10) we find

$$-\frac{dP_1}{dt} = \frac{\pi a^4}{16\eta l V}(P_1^2 - P_2^2)$$

or

$$\begin{aligned} t &= \frac{16\eta l V}{\pi a^4}\int_{P_4}^{P_3}\frac{dP_1}{P_1^2 - P_2^2} \\ &= \frac{8\eta l V}{P_2 \pi a^4}\int_{P_4}^{P_3} dP_1 \left(\frac{1}{P_1 - P_2} - \frac{1}{P_1 + P_2}\right) \\ &= \frac{8\eta l V}{P_2 \pi a^4}\log\left(\frac{P_3 - P_2}{P_4 - P_2}\cdot\frac{P_4 + P_2}{P_3 + P_2}\right) \quad \text{..........} \quad (11) \end{aligned}$$

Transition between Poiseuille flow and Knudsen flow.—If the pressure in a gas is gradually reduced, viscosity ceases to have any meaning when the mean free path becomes comparable with the radius of the tube. On p. 131 we showed that, at very low pressures, F was proportional to a^3 rather than a^4, and we gave plausible reasons for thinking that, for the intermediate region, formula (9) might be corrected by a factor $(1 + Cl/a)$. We can also approach this question from the hydrodynamic point of view thus: we assume that the molecules lose all their drift velocity on striking the wall, but are freely accelerated by the pressure-gradient in a region extending inward from the wall for a distance l that is, until a molecule that has left the wall has made, on an average, one collision. In the remaining region we assume that the distribution of velocities is controlled by viscosity in the ordinary way. The argument used in deriving equation (6) still applies, with the difference that the constants A and B are chosen, not so that v vanishes at the walls, but so that, at $r = a - l$, it agrees with the value v_1 which we derive in the following way. If the mean velocity of an individual molecule perpendicular to the axis of the tube is $\bar{c}_r$, it will collide with the walls $\bar{c}_r/2a$ times/sec., each time losing the average drift velocity v_1. Since the number of molecules hitting the walls per second is not affected by collisions, the total momentum transferred to the walls per second by unit volume of gas is $Nm\,\dfrac{\bar{c}_r}{2a}\,.\,v_1$ so that the momentum transferred per second to a length of wall dx is $\dfrac{Nmv_1\bar{c}_r}{2a}\,\pi a^2 dx$, and this is the retarding drag acting on all the gas, which has to be equal to the net force tending to push the gas forwards, that is, to $-\,dx\,.\,\dfrac{\partial P}{\partial x}\,.\,\pi a^2$, so that

$$v_1 = -\frac{2a}{\rho\bar{c}_r}\frac{\partial P}{\partial x} \quad \dots\dots\dots\dots\dots\dots \quad (12)$$

On p. 148 we showed that, for laminar flow of viscous fluid in a tube, the velocity was given by the function

$$v = \frac{Ar^2}{4} + B \quad \dots\dots\dots\dots\dots\dots \quad (13)$$

Substituting this in equation [5(a)] we find that $A = \dfrac{1}{\eta}\dfrac{\partial P}{\partial x}$, but B has now to be determined so that at $r = a - l$, $v = v_1$ instead of v vanishing at $r = a$, as it does for a liquid.

Substituting these values, we obtain for the distribution of velocities

$$v = v_1 + \frac{r^2}{4\eta}\frac{\partial P}{\partial x} - \frac{(a-l)^2}{4\eta}\frac{\partial P}{\partial x}.$$

Substituting for v_1 and using the relation $\eta = \frac{1}{3}\rho \bar{c} l$ (p. 125) and neglecting the difference between $\bar{c}_r$ and $\bar{c}$, we find

$$v = \frac{\partial P}{\partial x}\left(-\frac{2al}{3\eta} - \frac{(a-l)^2}{4\eta} + \frac{r^2}{4\eta}\right)$$

$$= -\frac{\partial P}{\partial x}\frac{1}{4\eta}\left(\left\{a + \frac{l}{3}\right\}^2 - r^2\right) \quad \ldots\ldots\ldots\ldots \quad (14)$$

neglecting terms of order l^2.

Comparison of this with expressions (6) and (7) shows that the gas behaves as if it were an ordinary viscous fluid flowing without slip in a tube of radius $(a + l/3)$. To get the relation between total flux and pressure gradient, corresponding to equation (9), we must integrate $2\pi r v dr$ between $r = 0$ and $r = a - l$, and add to it the flux through the ring between $r = a - l$ and $r = a$, where we may suppose that the velocity is equal to v_1, as, on the average, a molecule travels a distance l between its last collision with another molecule and its impact with the wall. Therefore, on an average, the drift velocity is v_1 up to a distance l from the wall. At greater distances, the effect of mutual collisions between molecules becomes important, which may be described by means of a coefficient of viscosity.

This gives

$$F = -\frac{2\pi}{4\eta}\frac{\partial P}{\partial x}\left(\left\{a^2 + \frac{2al}{3}\right\}\frac{(a-l)^2}{2} - \frac{(a-l)^4}{4}\right) - \frac{2al}{3\eta}\frac{\partial P}{\partial x}2\pi a l$$

$$= -\frac{\pi}{8\eta}\frac{\partial P}{\partial x}\left\{a^4 + \frac{4a^3 l}{3}\right\} \simeq -\frac{\pi}{8\eta}\frac{\partial P}{\partial x}\left(a + \frac{l}{3}\right)^4 \quad \ldots\ldots\ldots\ldots \quad (15)$$

neglecting terms of order l^2. This is again consistent with what would be obtained for an ordinary fluid flowing in a tube of radius $a + l/3$.

Thus we have derived, by another route, the result inferred on p. 131, namely, that the form of correction of the laws of viscous flow when l is becoming appreciable is a factor of the form $1 + Cl/a$. The above argument shows us a little more than the former one, namely, that the effect of a finite l is to *increase* the rate of flow above the "classical" value. This point was left undecided by the former argument.

The transition between Poiseuille flow and turbulent flow.—For a liquid, the above equations represent *possible* states of flow in a tube (at least as long as the stream velocity is small compared with the velocity of sound in the liquid, so that we can neglect the compressibility), but, in practice, laminar flow breaks down long before this.

This occurs because the solution of the hydrodynamic equations is not *uniquely* fixed by the boundary conditions. For a given state of motion of the boundary walls there are, in fact, infinitely many possible states of motion of the fluid itself, and, for an actual state of motion to be realised, it is necessary, not only that it should satisfy the equations of motion, but also that it should be *stable*, at least against small disturbances. A state of affairs of this kind is very common in physics, although the opposite situation, that the state of affairs is uniquely prescribed by the differential equations and the boundary conditions, is equally common (for example, the electrostatic field within an enclosure is completely determined if the normal component is known over a closed surface surrounding it).

We are therefore led to examine two questions: (*a*) In what conditions is a state of laminar flow stable? (*b*) What is the nature of the motion that replaces it when the velocities become too high?

Question (*a*) has proved extremely intractable, though it has been studied theoretically and experimentally since the time of **Kelvin**, and it is only in one or two cases that definite results have been obtained. **G. I. Taylor** studied the motion of a liquid between two concentric rotating cylinders and was able to predict theoretically, and to verify experimentally, the stage at which the low-velocity regime (the lines of flow being circles surrounding the axis of the cylinders) should break down. He found that this regime was replaced by an elaborate structure of *vortices*, arranged between the cylinders somewhat in the manner of "roller-bearings". This may be regarded as the first stage in the transition between laminar motion and turbulence. Two more stages probably occur as the velocities become still higher, the vortices probably become of smaller dimensions and more numerous and finally begin to dash about the liquid in random manner. This is one of the most characteristic features of ordinary "turbulence" in a pipe, in contrast with the vortices found by Taylor which are practically stationary. Broadly speaking, a similar set of occurrences takes place behind a solid obstacle in a steadily-flowing liquid. At small velocities, there is evidence of eddying motion in the "dead" liquid immediately behind the solid [Fig. 54(*a*)], as the velocity increases this becomes more pronounced, until one reaches a state of affairs in which vortices are detached alternately from the right- and left-hand side of the obstacle and form a "wake" with large isolated vortices spaced [Fig. 54(*b*)] in the manner of lamp posts on the opposite sides of a street. This formation is known as the "Karman vortex-street", and is sometimes visible when a flat body, such as the paddle of a canoe, is pushed through the water with moderate speed. As the speed increases, these isolated vortices disappear, and the final state of affairs is just an eddying "wake" extending for long distances behind the obstacle [Fig. 54(*c*)].

The prediction of the critical velocities at which these various changes take place is almost impossible in any actual case, yet we can use dimensional arguments (Chapter X) to make predictions about the way in which a critical velocity which has dimensions $[LT^{-1}]$ might be expected to vary with a (the relevant dimension of the apparatus, *e.g.* radius of tube, breadth of obstacle), and with η the viscosity and ρ the density. The dimensions of these quantities are

$$v_c = [L]\,[T]^{-1}, \quad a = [L], \quad \eta = [M]\,[L]^{-1}\,[T]^{-1}, \quad \rho = [M]\,[L]^{-3}.$$

By carrying out the process of comparison of indices explained in Chapter X, we can easily show that the only non-dimensional quantity that can be formed out of these four is the Reynolds number $R = v_c a\rho/\eta$, which also appears in Chapter X in the slightly different context of determining the *force* on an obstacle. Our conclusion is that, for a given geometry of apparatus, turbulence should set in at a constant value of R, and this prediction is well borne out in practice, whether we vary a or η. In flow through a tube, the critical value of R is of the order of thousands, but is sensitive to the exact conditions at the entrance and exit.

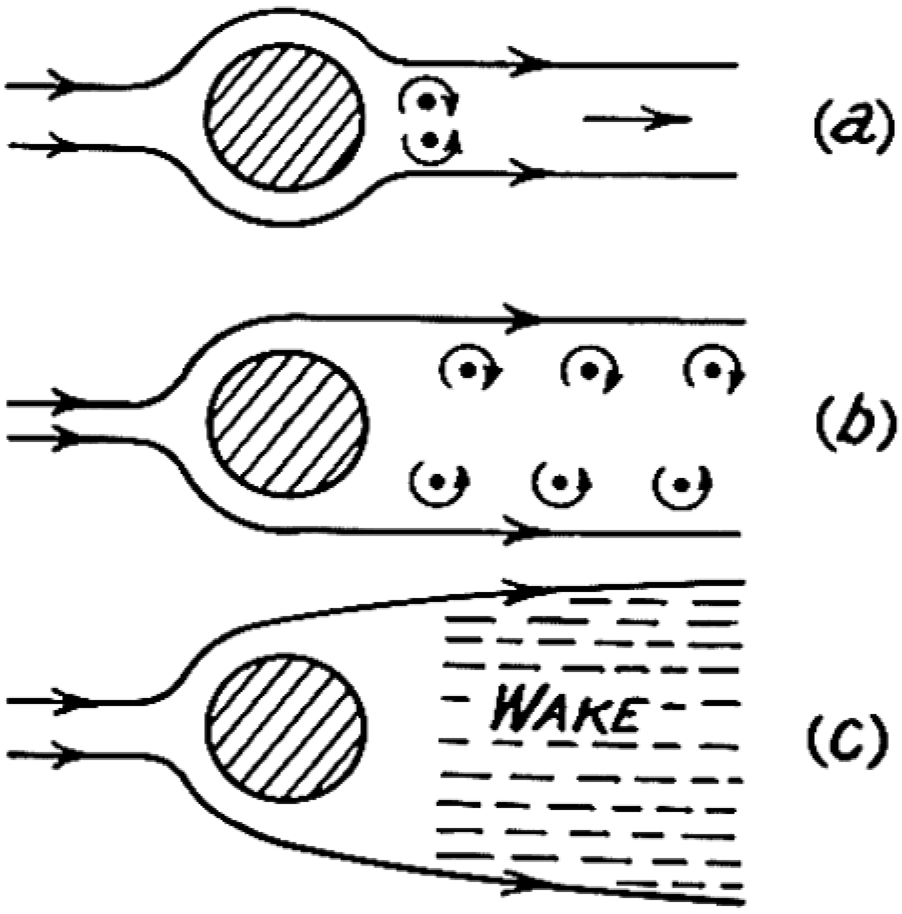


FIG. 54.

We have already mentioned some of the characteristics of turbulent flow. The most obvious one is the much more thorough mixing of the fluid that takes place, owing to the existence of moving vortices. This fact is of great importance in many fields, ranging from meteorology to chemical engineering. For the same reason the rate of flow of liquid in a pipe is practically uniform over the cross-section (the effective component of drift velocity being practically constant except in the immediate neighbourhood of the walls) in marked contrast with the parabolic law characteristic of Poiseuille flow. Perhaps the most characteristic property of all is the change in the relationship between velocities and forces acting on or exerted by the fluid. For laminar flow, this relationship is a *linear* one, steady velocities being proportional to forces or pressure gradients, as for Poiseuille flow that we have just studied, or for **Stokes'** law giving the force on a sphere of radius a moving very slowly through a fluid with velocity v as $F = 6\pi\eta av$. (The reader is referred to Lamb's *Hydrodynamics* for a discussion of this problem.) For turbulent flow, *mean velocities are very nearly proportional to the square roots of forces or pressure-gradients.* This applies both to flow in pipes, for which many empirical formulae have been devised, and to the motion of solids through fluids. For the latter, the change-over from a Stokes' type force to a square-law force usually takes place at a Reynolds number of about 20. Once we are well above the transition region, the force is proportional to the square of the velocity, and becomes *independent of the viscosity.* We may regard a viscosity as "small" in the sense described on p. 140 above, when it no longer affects the resistance. This fact makes possible a simple interpretation, due to **Newton**, of this type of resistance. One may think of the wake as being practically at rest, and the remaining liquid as represented by a wide jet of density ρ and velocity v, incident on the body supposed held at rest. If all the momentum of the liquid that actually hits the body is destroyed, the net force on the body would be $\rho Av.v$, where A is the greatest area of cross-section. For a "rounded-off" shape, such as a sphere or paraboloid, only a fraction of this momentum is destroyed, twice this fraction being known as the *drag coefficient*, so that

$$F = \tfrac{1}{2}C_D\rho Av^2 \quad\ldots\ldots\ldots\ldots\ldots\ldots\ldots\quad (16)$$

The quantity C_D is practically independent of the Reynolds number, but depends mainly on the shape of the body. For a sphere, it is found experimentally to be about 0·3 for a very wide range of velocities, but for a very flat body it may rise above unity. This arises from the fact that, besides the destruction of momentum at the front of the body, eddies in the wake behind the body may reduce the pressure *behind* the body by the Bernoulli effect [which is also proportional to $\rho\times$ the square of the velocity (p. 145)].

5. Irrotational motion

We have, in Art. 4, studied some particular solutions of the equations of hydrodynamics, without enquiring very closely whether or not other solutions satisfying the same boundary conditions were also possible. We have, in fact, found cases where more than one type of solution may be physically relevant as well as mathematically possible. It is therefore natural to ask what restrictions are needed, as well as the boundary conditions, to make the solution of a particular problem *unique*. If the viscosity is supposed to be zero, these conditions take a very simple form, namely, that the motion must be *irrotational*, that is the velocity vector must everywhere satisfy the equations

$$\frac{\partial v_z}{\partial y} - \frac{\partial v_y}{\partial z} = \frac{\partial v_x}{\partial z} - \frac{\partial v_z}{\partial x} = \frac{\partial v_y}{\partial x} - \frac{\partial v_x}{\partial y} = 0, \quad \text{or curl } \bar{v} = 0 \; . \; . \quad (17)$$

It is probably true to say that these conditions are never satisfied exactly. In general, they imply tangential slipping at solid-liquid or liquid-liquid boundaries which is almost certainly physically inadmissible. Why, then, has so much effort been put into the study of irrotational motion? There are probably three reasons:

(*a*) Much of the mathematics can be applied, *mutatis mutandis* to other branches of physics such as gravitation and electrostatics, so did not have to be specially invented for this purpose.

(*b*) There are problems such as the motion of waves in an open sea, where the equations (17) are probably true to a high approximation.

(*c*) For other problems, where equations (17) are not satisfied, determination of the irrotational solution forms a foundation on which a physical theory can be gradually erected.

We now proceed to examine some irrotational solutions of the equations of hydrodynamics, fully realising that, by themselves, they are liable to lead to physically absurd conclusions (for example, that there is no force on a body moving steadily through a liquid). We introduce a quantity known as the *velocity potential*. For velocities satisfying equation (17), it is always possible to find a quantity ϕ such that

$$v_x = -\frac{\partial \phi}{\partial x}, \quad v_y = -\frac{\partial \phi}{\partial y}, \quad v_z = -\frac{\partial \phi}{\partial z} \quad \text{or } \bar{v} = -\,\text{grad}\,\phi \; . \; . \; . \quad (18)$$

For proof that *all* solutions of equation (17) satisfy equation (18), the reader is referred to textbooks of vector analysis. [The converse can be proved quite easily by direct substitution of v_x, etc., from equation (18) into equation (17).] The velocity potential is a distinctly artificial conception, and has no direct physical interpretation. It is thus similar to quantities like the vector potential in electromagnetics, or the wave-function in wave-mechanics, which are introduced solely to enable us to calculate other quantities, and cannot be experimentally measured themselves. Equation (18) shows that the lines of flow in irrotational motion are perpendicular to the equipotentials, or lines of constant ϕ.

A. Acoustic type solutions

We use equations (1) and (2) (p. 142), but neglect all powers and products of the components of $\bar{v}$ and of all quantities, such as the derivatives of ρ, which would also vanish if the velocities vanished. We also assume a simplified equation of state $P = (\rho - \rho_0)c^2$, where c is the velocity of sound and $c^2 = \partial P/\partial \rho$. We also neglect any external field of force.

Equation (1) becomes

$$\rho\left(\frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} + \frac{\partial v_z}{\partial z}\right) = \frac{\partial \rho}{\partial t} = \frac{1}{c^2}\,\frac{\partial P}{\partial t} \quad \ldots\ldots\ldots\ldots \quad (19)$$

or

$$\rho \operatorname{div} \bar{v} = \frac{1}{c^2}\,\frac{\partial P}{\partial t}.$$

Equation (2) becomes

$$\frac{\partial v_x}{\partial t} = -\frac{1}{\rho}\,\frac{\partial P}{\partial x} \text{ and two similar equations,}$$

or

$$\frac{\partial \bar{v}}{\partial t} = -\frac{1}{\rho} \operatorname{grad} P \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (20)$$

Inserting values from equation (18), equation (20) becomes

$$-\frac{\partial^2 \phi}{\partial x . \partial t} = -\frac{1}{\rho}\,\frac{\partial P}{\partial x},$$

and two similar equations, or

$$\frac{\partial}{\partial t}\,(\operatorname{grad} \phi) = -\frac{1}{\rho} \operatorname{grad} P.$$

This integrates into

$$P = \rho\,\frac{\partial \phi}{\partial t} + F(t) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (21)$$

where F is independent of x, y, and z. [A finite value of $F(t)$ would correspond to the application of a uniform pressure throughout the whole of the liquid.]

Substituting in equation (19) we find finally

$$\frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} + \frac{\partial^2 \phi}{\partial z^2} - \frac{1}{c^2}\,\frac{\partial^2 \phi}{\partial t^2} = 0.$$

or

$$\nabla^2 \phi - \frac{1}{c^2}\,\frac{\partial^2 \phi}{\partial t^2} = 0 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (22)$$

This is the wave-equation, and, once a suitable solution is found, $\bar{v}$ can be calculated by means of equation (18) and P by equation (21). For further discussion, the reader is referred to textbooks of sound, where the wave equation often contains S, the condensation or fractional change of density, instead of ϕ. S is directly proportional to excess pressure. By differentiating equation (22) we can show that P, and each of the three components of $\bar{v}$, themselves satisfy the wave equation, so that the use of the velocity potential leads here to no particular advantage.

B. Incompressible type solutions

Here we assume that ρ is constant, so that we do not need an equation of state connecting P and ρ, but we retain the non-linear terms in the hydrodynamic equation themselves. Equation (1) becomes, after using equation (18),

$$\frac{\partial^2\phi}{\partial x^2}+\frac{\partial^2\phi}{\partial y^2}+\frac{\partial^2\phi}{\partial z^2}=0 \quad \text{or} \quad \nabla^2\phi=0 \quad \ldots\ldots\ldots \quad [22(a)]$$

Equation (2) becomes, after using equation (17), to turn $v_y \dfrac{\partial v_x}{\partial y}$ into $v_y \dfrac{\partial v_y}{\partial x}$, etc.

$$-\rho\frac{\partial^2\phi}{\partial x.\partial t}+\rho v_x\frac{\partial v_x}{\partial x}+\rho v_y\frac{\partial v_y}{\partial x}+\rho v_z\frac{\partial v_z}{\partial x}=-\frac{\partial P}{\partial x}+F_x$$

and two similar equations,

or $$-\rho\frac{\partial}{\partial t}\operatorname{grad}\phi+\frac{\rho}{2}\operatorname{grad}v^2=-\operatorname{grad}P;\ \text{if } F_x,\ F_y,\ F_z=0.$$

Integrating $$P=G(t)+\rho\frac{\partial\phi}{\partial t}-\frac{\rho}{2}v^2 \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (23)$$

where $G(t)$ is independent of the coordinates, and corresponds to a uniform application of pressure. If an external force, such as gravity, which can be associated with a potential energy Ω per unit volume of fluid, is also acting, we must write $F_x=-\partial\Omega/\partial x$, $F_y=-\partial\Omega/\partial y$, etc., and equation (23) must be replaced by

$$P=G(t)+\rho\frac{\partial\phi}{\partial t}-\frac{\rho}{2}v^2-\Omega \quad \ldots\ldots\ldots\ldots \quad [23(a)]$$

where the constant of integration for Ω can be absorbed into the arbitrary function $G(t)$.

Equation [22(a)] is the same as that satisfied by the electrostatic potential in free space, and therefore we can take over the mathematical theory as it stands, and use, in the new physical context, any results for which we can find a hydrodynamical interpretation. The following are well-known theorems in potential theory. The reader who desires formal proofs of them is referred to Jeans, *Electricity and Magnetism*, Lamb's *Hydrodynamics*, or to standard treatises on potential theory.

(i) The distribution of electrostatic potential within a volume containing no free charges is uniquely determined if the potential is known at every point on a closed surface surrounding that volume.

(ii) The distribution of electrostatic potential within a volume containing no free charges is uniquely determined if the normal component of field is known at every point on a closed surface surrounding that volume.

(iii) Let f be *any* function satisfying a boundary condition of the type (i) or (ii). Then the integral of $(\partial f/\partial x)^2 + (\partial f/\partial y)^2 + (\partial f/\partial z)^2$ or $(\text{grad } f)^2$ over the volume, takes a minimum value if f also satisfies Laplace's equation.

To translate these into hydrodynamic form, we replace electrostatic potentials by velocity potentials, and the three components of electric field by the corresponding components of velocity. Then, for example, theorem (ii) is to be replaced by the statement that *associated with any given motion of the boundaries surrounding an incompressible non-viscous liquid, there is only one possible state of* irrotational *motion. In particular, if all the boundaries are at rest, no irrotational motion is possible.* A generalisation, due to Kelvin, of theorem (iii), further shows that, *given any set of boundary conditions, of all the possible states of liquid motion consistent with them, the irrotational one is that with the lowest kinetic energy.* Here the correspondence is between electrostatic energy (proportional to the square of the electric field vector) and kinetic energy (proportional to the square of the velocity vector).

Naturally, the analogy between electrostatics and irrotational motion cannot be pushed too far. There is, for example, no quantity in electrostatics corresponding to pressure, nor are there any hydrodynamic quantities corresponding to the dielectric constant or dielectric displacement. As an example of a hydrodynamic theorem that has no precise counterpart in electrostatics we may mention:—

(iv) *In a non-viscous incompressible fluid the circulation* [*defined as a vector whose x-component is the integral of* $\partial v_z/\partial y - \partial v_y/\partial z$, *or* $(\text{curl } v)_x$, *over the whole volume of the fluid*] *is constant and cannot be altered by any motion of the boundaries of the fluid.*

In practice, these irrotational solutions are hardly ever realised, even in a liquid of very small viscosity, because they imply (in general) slipping of liquid along solid boundaries. This arises because we cannot at the same time assign prescribed values to both normal and tangential components of an electric field all over any boundary surface.

Similarly, we cannot at the same time prescribe the normal and tangential components of velocity over the boundaries of a fluid moving irrotationally, *e.g.* if a portion of the boundary surface is fixed, it is always possible to make the velocity normal to it zero, but then the tangential component is completely determined, and is, in general, *not* zero. What happens in practice is that these regions, where, according to irrotational theory, slipping should occur, act as *sources for the generation of vortices*, which then move outwards into the remaining liquid, this process being governed by laws very similar to those of heat conduction, and are finally dissipated by viscosity. Recently, this process of formation and diffusion of vorticity has been extensively studied theoretically and experimentally, and the regions in question are usually known as *boundary layers*. A simple example of a boundary layer is the region immediately behind an obstacle in a moving stream, already described above (p. 153).

Types of boundary condition. Free and rigid boundaries

We have already stated the boundary condition that must hold at a rigid boundary surface, namely, that the normal component of the velocity in the fluid must be the same as that of the rigid surface. Sometimes, however, we have to consider conditions at a free surface. An example of such a problem is the determination of the form of a steadily revolving liquid (p. 146). Here the boundary condition is that the pressure is constant all over the free surface. Equation (22) shows that this implies $\partial\phi/\partial t = 0$ (provided that the stream velocity is so small that its square may be neglected). This means that, everywhere on the free surface, ϕ remains constant in time, which is thus the boundary condition appropriate to a free surface. In practice, this usually means that ϕ takes the same value all along the free surface, but this is slightly more stringent than is really required by equation (22). The condition that ϕ is constant has an electrostatic analogy, for we know that electrostatic potential is constant on the surface of a *conductor*, so that we may think of the boundary of a conductor as analogous to a free surface. The analogy with a rigid surface is less obvious. Corresponding to the vanishing of the normal component of the velocity, the normal component of the electric field must vanish, and this is what might be expected in an electrostatic problem if the bounding medium were one of zero dielectric constant.

It is now possible to build up the theory of irrotational motion in complete analogy with that of electrostatics, the method being the selection of suitable solutions of Laplace's equation to satisfy as many of the conditions of the problem as possible, the effect of free and rigid surfaces being allowed for, *e.g.* by the method of images. Other specialised methods used in electrostatics such as inversion and, for two-dimensional problems, the use of complex variable theory, are also applicable. For a full account of such methods, the reader is referred to Milne-Thomson's *Theoretical Hydrodynamics* or Lamb's *Hydrodynamics.* We shall mention only a few typical problems here.

(i) *The isolated source.*—If we write, using spherical polar coordinates, $\phi = A(t)/r$, this is a solution of Laplace's equation, corresponding to an isolated electric charge in electrostatics. Differentiating with respect to r, we have

$$v_r = \frac{A(t)}{r^2} \qquad \text{(24)}$$

while, from equation (23), the pressure is given by

$$\frac{P}{\rho} - gz = \frac{\partial\phi}{\partial t} - \tfrac{1}{2}v_r{}^2 = \frac{1}{r}\frac{dA}{dt} - \frac{1}{2}\frac{A^2}{r^4} \qquad \text{(25)}$$

(if gravity is taken into account).

This solution was applied by **Lamb** in an attempt to describe under-water explosions. If an under-water explosion generates a bubble of radius a, equation (24) shows that the velocity of the water will agree with the outward velocity da/dt of the bubble surface if we put $A(t) = a^2 \dfrac{da}{dt}$. Substituting this result in equation (25), we find, putting $r = a$ and supposing that the pressure of the gas inside the bubble is a function of its radius only,

$$\frac{P(a)}{\rho} - gz = a\frac{d^2a}{dt^2} + \frac{3}{2}\left(\frac{da}{dt}\right)^2 \qquad \text{(26)}$$

This equation for a can be solved numerically if $P(a)$ be assumed known, that is, if we suppose the pressure to be a function of the volume only. This is only a first approximation to a theory of an under-water explosion, and it is now known that the pressures are so high that the compressibility of the water, and the finite time of propagation of pressure differences in the gas-bubble itself, have to be considered. The inclusion of the term gz, representing the ordinary hydrostatic pressure in the water, has an interesting consequence. However large $P(a)$ is initially, there must come a time when the pressure $P(a)$ in the bubble falls below that proper to the depth of water outside, so that the outward acceleration eventually becomes negative and the bubble eventually stops expanding and contracts again, the motion being *oscillatory.* This type of motion was predicted theoretically by **Butterworth,** and has been experimentally confirmed, a typical solution being shown in Fig. 55.

It will be seen from Fig. 55 that the minima are very much sharper than the maxima. To see the reason for this, multiply equation (25) on both sides by $4\pi a^2 \frac{da}{dt}$, or dV/dt, where $V = 4\pi a^3/3$ is the volume of the bubble. Then

$$(p - \rho g z)\frac{dV}{dt} = \frac{d}{dt}\left[2\pi\rho a^3\left(\frac{da}{dt}\right)^2\right] \quad \ldots\ldots\ldots\ldots (27)$$

The left-hand side of equation (27) represents the work being done by the pressure of gas in the bubble pushing the surface outwards, so that the right-hand side must represent the rate of change of kinetic energy of the water. Thus, the bubble behaves mathematically as if it were a mass $4\pi\rho a^3$ moving with a velocity da/dt, and the motion is formally similar to that of a mass on a spring, if one supposes the mass to vary during the oscillation. The sharpness of the minima in Fig. 55 is thus accounted for since the effective mass is very small when a is small, and the "effective frequency" of the oscillation is correspondingly much higher at this epoch than it is near the maxima.

We have been led to describe the kinetic energy of the water by means of a hypothetical mass which is supposed to move with the velocity da/dt of the boundary surface. A similar device can be used to allow for the fact that, when a solid moves through a liquid, kinetic energy is imparted to the liquid. This can also be described, quite apart from any *resistance* offered by the water, by the apparent addition of mass to the true mass of the solid. This extra mass, which is usually of the order of magnitude of that of the liquid displaced by the solid, is often called "hydrodynamic virtual mass". An example is given on p. 164.

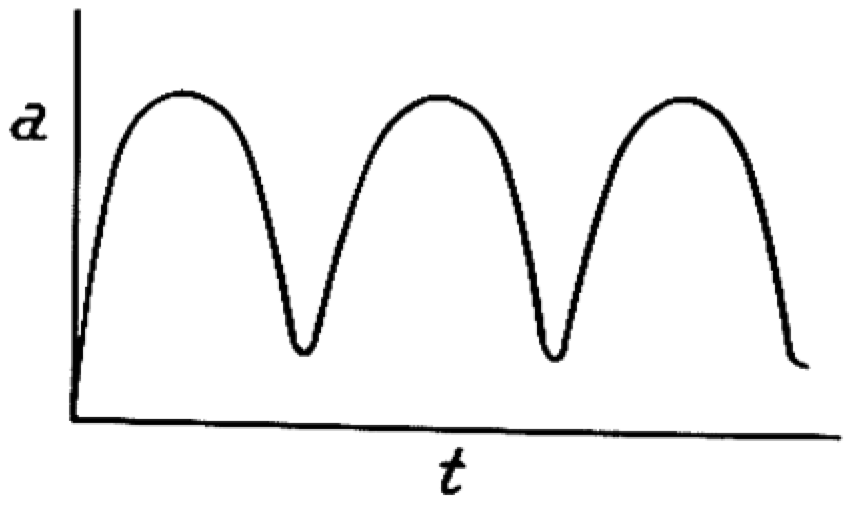


FIG. 55.

(ii) *The motion of a sphere through a perfect fluid.*—This problem is the hydrodynamic counterpart of the electrostatic problem of the disturbance of a uniform electric field by the presence of a conducting sphere. It is convenient to work with coordinates in which the sphere is at rest, and the fluid moving past it with a velocity U. We have to find a velocity potential ϕ with the following properties:—

1. $\partial\phi/\partial r$ must vanish at $r = a$, the surface of the sphere. This corresponds to the fact that the component of velocity normal to a solid surface must vanish.

2. At a great distance from the sphere ϕ must correspond to uniform motion with velocity U, which we may suppose to be parallel to the z-axis. This is

$$\phi_1 = -Uz = -Ur\cos\theta$$

in polar coordinates. We now correct ϕ_1, to allow for the presence of the sphere, by adding a term ϕ_2 which vanishes for r large. ϕ_2 must also satisfy Laplace's equation, and be symmetrical about the z-axis. Such functions are known as *axial harmonics*, and it is known that the most general possible form for a function satisfying these conditions is (see Jeans, *Electricity and Magnetism*)

$$\phi_2 = \frac{A}{r} + \frac{B\cos\theta}{r^2} + \frac{C(3\cos^2\theta - 1)}{r^3} + \ldots$$

To satisfy the condition $\partial\phi/\partial r = 0$ at $r = a$, we must have

$$0 = \left(\frac{\partial\phi_1}{\partial r}\right)_{r=a} + \left(\frac{\partial\phi_2}{\partial r}\right)_{r=a} = -U\cos\theta - \left(\frac{A}{r^2} + \frac{2B\cos\theta}{r^3} + \ldots\right)_{r=a}$$

which can be satisfied if we put $B = -\frac{1}{2}Ua^3$, $A = C = D = \ldots = 0$. It can be shown further that no other function ϕ_2 can be found to satisfy all these conditions (except one differing from ϕ_2 by an additive constant only), but the "uniqueness theorems" required for this are beyond the scope of this book, and the reader is referred to treatises on potential theory. The requirement is that $\partial\phi/\partial r$ should vanish at $r = a$ for all values of θ and, since ϕ_2 contains a term that *varies with θ in the same way as* ϕ_1, it is on common-sense grounds redundant to include any others. (If ϕ_1 varied in a more complicated way with θ, it would be necessary to include more terms of the series, possibly even an infinite number of such terms.)

Our complete solution, representing the motion of fluid past a fixed sphere of radius a, is thus

$$\phi = -Ur\cos\theta - \frac{1}{2}\frac{Ua^3\cos\theta}{r^2} \quad \ldots\ldots\ldots\ldots \quad (28)$$

We can change to coordinates *fixed relative to the fluid* by superimposing a velocity $-U$ upon everything, which involves adding the quantity $Ur\cos\theta$ to ϕ, and we conclude that the velocity potential associated with the motion of a sphere through a stationary fluid with velocity U is

$$\phi' = -\frac{1}{2}\frac{Ua^3\cos\theta}{r^2} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (29)$$

By differentiating this expression, we can determine the components of velocity at each point and trace out the lines of flow, which form the same pattern as the lines of force of a small magnet or electric dipole

(Fig. 56). If we compute the lines of flow corresponding to the velocity potential (28) for the motion of fluid with a velocity U past a stationary sphere we arrive at Fig. 57, which is the same pattern as that produced by a small magnet or dipole in a uniform field, expression (29) being just the electrostatic or magnetic potential associated with a dipole of moment $\frac{1}{2}Ua^3$. Thus, corresponding to the result that the effect of a conducting sphere in disturbing a uniform field can be represented by means of a dipole of suitable strength at its centre, we have the result that the disturbance of uniform flow by a solid sphere can be represented by means of a "double source", *i.e.* a combination of a source and sink of equal strengths analogous to an electric or magnetic dipole.

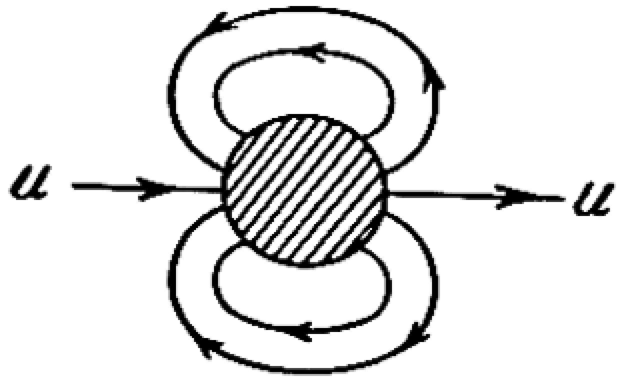


FIG. 56.

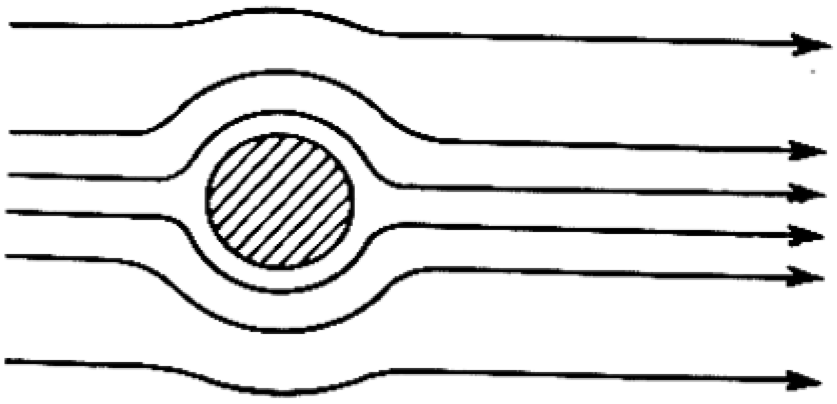


FIG. 57.

We shall now calculate the total kinetic energy of the fluid corresponding to the velocity potential (29). The component of velocity along r is

$$-\left(\frac{\partial\phi'}{\partial r}\right)=\frac{Ua^3\cos\theta}{r^3},$$

while the component perpendicular to r is

$$-\frac{1}{r}\left(\frac{\partial\phi'}{\partial\theta}\right)=\frac{1}{2}\,\frac{Ua^3\sin\theta}{r^3}.$$

The kinetic energy of a small volume dV of fluid with coordinates r, θ is then

$$\frac{\rho dV}{2}\left[\left(\frac{\partial\phi'}{\partial r}\right)^2+\frac{1}{r^2}\left(\frac{\partial\phi'}{\partial\theta}\right)^2\right]=\frac{\rho dV U^2a^6}{2r^6}\left(\cos^2\theta+\frac{1}{4}\sin^2\theta\right).$$

Fig. 58 shows that the area lying between r and $r+dr$ and θ and $\theta+d\theta$ is $rdr.d\theta$, and since we have symmetry about the z or θ axis, the volume dV lying between these limits of r and θ is obtained by rotating the area $rdrd\theta$ about the z-axis, which is at a distance $r\sin\theta$, so that

$$dV=2\pi r^2dr\sin\theta d\theta.$$

The kinetic energy of the liquid in the volume dV is thus equal to

$$dT = \frac{\pi\rho U^2 a^6}{r^4}\, dr \sin\theta d\theta \left(\cos^2\theta + \tfrac{1}{4}\sin^2\theta\right),$$

which must be integrated over the whole region outside the sphere, that is from $r = a$ to infinity and from $\theta = 0$ to π, so that

$$T = \pi\rho U^2 a^6 \int_a^\infty \frac{dr}{r^4} \int_0^\pi \left(\cos^2\theta + \tfrac{1}{4}\sin^2\theta\right)\sin\theta d\theta$$

$$= \frac{\pi\rho a^3 U^2}{3} = \frac{2\pi\rho a^3}{3} \cdot \frac{U^2}{2} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (30)$$

Thus, the total kinetic energy is the same as that of a body of mass $2\pi\rho a^3/3$ moving with a velocity U, so that the "hydrodynamic virtual mass" is *half the mass of the liquid displaced by the sphere.*

Comparison of Fig. 57 with Fig. 54 shows that the velocity potential (28) fails to describe the facts in one very important respect, namely, that the ideal solution shows complete symmetry in the flow-pattern upstream and downstream of the obstacle. A further consequence of this is that *the total force on the obstacle is zero for this type of flow.* To prove this, we substitute the velocity potential (28) in the equation (23) for the pressure. Under steady conditions $\partial\phi/\partial t$ is zero, and at $r = a$, the surface of the sphere v_r vanishes, so that we have

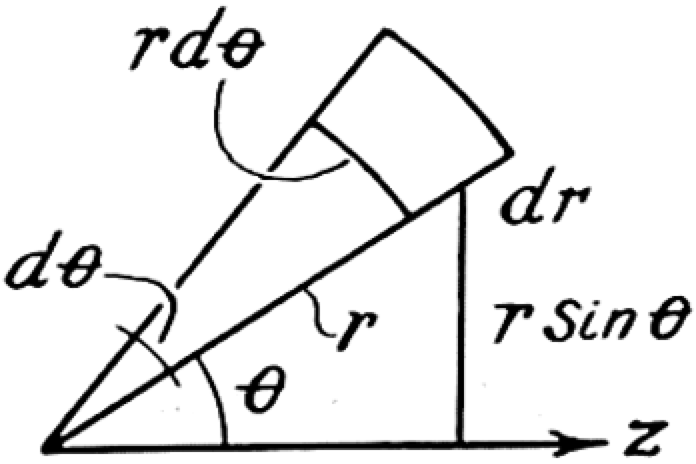


FIG. 58.

$$P = G(t) - \tfrac{1}{2}\rho v_\theta{}^2 = G(t) - \tfrac{1}{2}\rho\left[\frac{1}{r^2}\left(\frac{\partial\phi}{\partial\theta}\right)^2\right]_{r=a}$$

$$= G(t) - \tfrac{1}{2}\rho U^2[\sin\theta + \tfrac{1}{2}\sin\theta]^2 \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (31)$$

Now consider the zone of the sphere lying between θ and $\theta + d\theta$. Its area is $2\pi a^2 \sin\theta d\theta$, and the pressure P acts normally upon it. By symmetry, the total force on the sphere must act along the θ or z axis, the resolved parts of the force perpendicular to this axis cancelling for the whole ring. Thus, the total force on the ring is

$$2\pi P a^2 \sin\theta d\theta . \cos\theta$$

acting along the z-axis, and the total force on the whole sphere is

$$\int_0^\pi 2\pi P a^2 \sin\theta \cos\theta d\theta,$$

which vanishes when expression (31) is substituted for P. This result is quite general for any shape of obstacle; for irrotational flow the net force on *any* solid body is zero, so that this represents a serious limitation of "classical" hydrodynamics. In practice, the effect of viscosity can never be neglected in the region close to the sphere, because the solution (28) corresponds to slipping over the surface of the sphere, and this leads to rotational motion and the formation of a "wake" in the manner sketched above (p. 153).

The irrotational solution can, however, be regarded for some purposes as a reasonable first approximation to the truth, and it is often found that it can be used satisfactorily, *e.g.* for predicting pressure distributions over the *upstream* face of an obstacle. Thus, equation (31) indicates a change in pressure due to the insertion of the sphere from the value $G(t) - \frac{1}{2}\rho U^2$, which it would have in the absence of the sphere, to the value $G(t) - \frac{9}{8}\rho U^2 \sin^2\theta$, so that the effect of inserting the sphere is to increase the pressure at a point on it whose coordinate is θ by

$$\tfrac{1}{2}\rho U^2 (4 - 9 \sin^2\theta)/4,$$

that is, it is increased for regions for which $\sin\theta < \frac{2}{3}$, and decreased for the remaining regions. We already know that this cannot be correct for all values of θ, but it is found experimentally to be a fair approximation for the front face of the sphere $(0 < \theta < \pi/2)$.

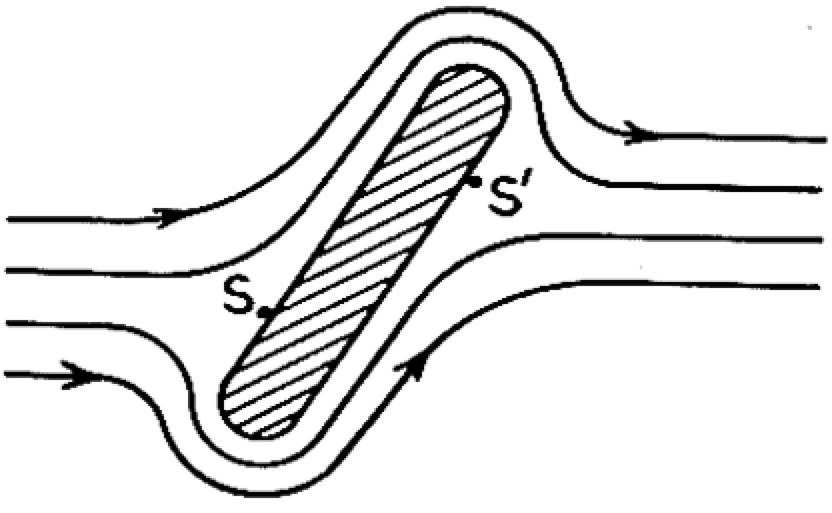


FIG. 59.

(iii) *The Rayleigh disc.*—Another problem that can be solved exactly, whose solution contains features of physical interest, is the flow of a perfect fluid past an ellipsoid. This problem can be solved even when the direction of flow is inclined to the axes of the ellipsoid. (The reader is referred to Lamb's *Hydrodynamics.*) While, in accordance with what we have just said, there is no net *force* as the result of such motion, there is, in general, a *couple* and its sign is such that it tends to turn the ellipsoid *broadside-on* to the stream. This is a slightly unexpected result—one might at first expect the stable regime to be the one causing least disturbance to the flow, the longest axis setting itself along the flow, but such is not the case. The lines of flow are sketched qualitatively in Fig. 59. S and S' represent the two *stagnation points* at which the velocity is zero, and these correspond to the neutral points in a magnetic field. Now, by Bernoulli's equation, the pressure is a maximum at a stagnation point, which makes it plausible that the couple should tend to turn the body broadside on. The exact calculations confirm this impression and show further that the couple is proportional to U^2, the square of the velocity, the constant of proportionality depending on the dimensions of the ellipsoid. The torque is thus independent of the *sense* of the flow, and it is this fact that was exploited by **Rayleigh** to make an instrument that could measure stream velocities in a sound-wave. His instrument consists essentially of a disc suspended on a torsion wire, the observed torque being proportional to the mean square of the stream velocity. The importance of the instrument depends on the fact that it is *absolute*, the disc can be regarded as an extreme case of an ellipsoid and the constant (torque/ρU^2) can be computed theoretically; also it is one of the few instruments available for a direct measurement of stream velocity. The irrotational solution is more satisfactory for *periodic* than for *steady* flow, because a wake is given no time to form (see p. 174).

It is well known to practical sailors that a boat cannot be kept on a straight course simply by lashing the helm in a central, or indeed in any other, position. The continual attention of the helmsman is needed to counteract the instability we have just mentioned, that tends to set the boat broadside-on.

(iv) *Waves on water.*—There are several well-defined types of wave on water and, according to the wavelength, the velocity of propagation may depend mainly on surface tension, in which case one often uses the term "ripple", on the depth of the water, in which case the term "long" or "canal" wave is often used, or on the wavelength itself, this last type usually being described as a "deep water wave". In all three types it is usually safe to neglect viscosity, and bring its effect in later on if it is desired to predict the rate of decay with distance or time. It is usually sufficient to assume irrotational motion, the effect of boundary layers in the neighbourhood of solids being of much less significance than it is for steady motion. The reason for this is simple, the stream velocity reverses itself periodically, so that the chances are that any vortices formed in the boundary layers will remain permanently near the solid boundaries where they are formed, rather than being carried bodily by the flow to other places in the fluid, as occurs with *steady* motion.

We shall restrict ourselves to plane waves, but shall give a general treatment that covers all three cases. We begin by asking in what circumstances a deformation of the free surface of a liquid into the form $\eta = \eta_0 \cos kx$ can be propagated with velocity V without change of form. Here η is the elevation of the free surface, x is measured in the direction of propagation of the wave, and y perpendicular to the free surface. We impose on everything a backward velocity V to bring the wave pattern to rest. Our problem is then to find a solution of Laplace's equation with lines of flow as shown in Fig. 60. On common-sense grounds we must conclude that the "waviness" of the lines of flow diminishes as we go deeper, until at the bottom (supposed flat) they become straight. We want to find the law governing this effect, also a relation between k or λ and V. (If λ is the wavelength, we have $k = 2\pi/\lambda$.)

The simplest type of velocity potential that we can assume, corresponding to an oscillatory motion of wavelength λ superimposed on a steady motion of velocity V, is

$$\phi = -Vx + Ae^{ky}\sin kx + Be^{-ky}\sin kx \quad \ldots\ldots\ldots\ldots (32)$$

where

$$k = 2\pi/\lambda.$$

It is easily verified that ϕ is a solution of Laplace's equation (22). We introduce also the auxiliary function

$$\psi = -Vy + Ae^{ky}\cos kx - Be^{-ky}\cos kx \quad \ldots\ldots\ldots (33)$$

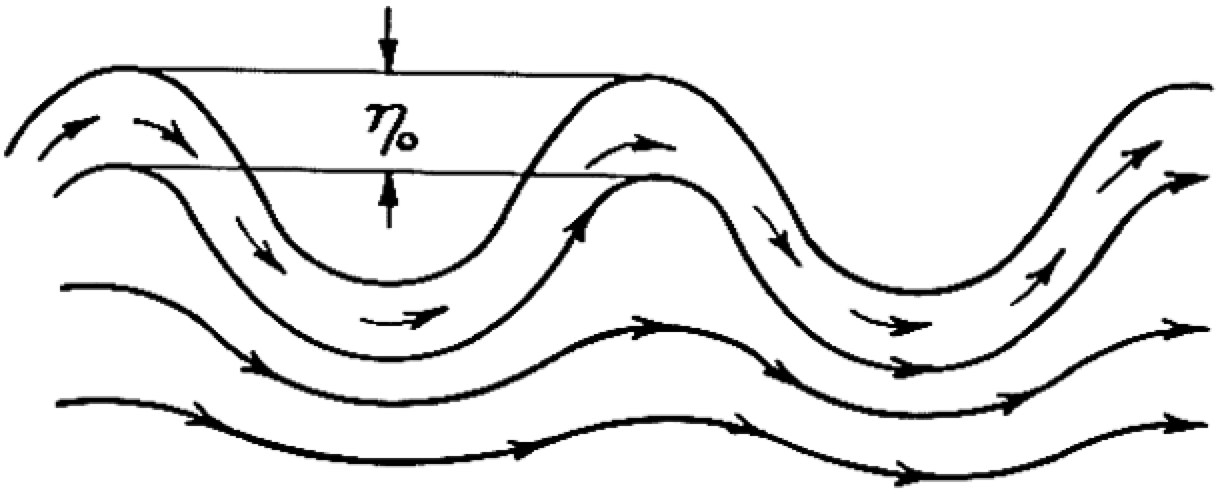


FIG. 60.

This function is sometimes known as the *stream function* and is *constant along lines of flow*, whereas the lines along which ϕ is constant are equipotentials which are perpendicular to the lines of flow (p. 155). To see this, we have from equations (32) and (33)

$$\frac{\partial\phi}{\partial x} = \frac{\partial\psi}{\partial y} \quad \text{and} \quad \frac{\partial\phi}{\partial y} = -\frac{\partial\psi}{\partial x} \quad \ldots\ldots\ldots\ldots\ldots (34)$$

so that the line ψ = constant is given by

$$0 = d\psi = \left(\frac{\partial\psi}{\partial x}\right)dx + \left(\frac{\partial\psi}{\partial y}\right)dy$$

or

$$\left(\frac{dy}{dx}\right)_{\psi\ \text{constant}} = -\frac{\left(\frac{\partial\psi}{\partial x}\right)}{\left(\frac{\partial\psi}{\partial y}\right)} = \frac{\left(\frac{\partial\phi}{\partial y}\right)}{\left(\frac{\partial\phi}{\partial x}\right)} = -\left(\frac{dx}{dy}\right)_{\phi\ \text{constant}} \quad \ldots (35)$$

so that the curves corresponding to the relations $\psi =$ constant and $\phi =$ constant are always at right angles. The student who is familiar with the theory of *conjugate functions* will recognise equations (34) as the Cauchy-Riemann equations, and ϕ and ψ are the real and imaginary parts of the analytic function $f(z) = -Vz + Ae^{ikz} + Be^{-ikz}$. He will also recognise equation (35) as a quite general property of conjugate functions.

We can now use the conditions at the surface and bottom to determine three relations between the four quantities A, B, V, and k. These are as follows:—

(*a*) *V has been chosen so that the wave pattern on the surface is stationary*, that is, the functions ϕ and ψ must represent motion in which the velocity is tangential to the free surface (Fig. 60), that is *the free surface is a streamline* and ψ is therefore constant along it. This implies from (33) that the equation to the free surface must be

$$\text{Constant} = -Vy + \cos kx\,(Ae^{ky} - Be^{-ky}) \quad \ldots\ldots\ (36)$$

We choose the origin of the y-coordinate to lie at the undisturbed level of the free surface. In this case ky never exceeds $k\eta_0$, where η_0 is the maximum elevation or depression of the free surface, and we make the assumption that η_0 is small compared with the wavelength, that is $k\eta_0 \leqslant 1$. It is then necessary to take the constant on the left-hand side of equation (36) zero, because $y = 0$ for undisturbed conditions. (We can always change either ϕ or ψ by adding constants to them without altering any physical results, because we are really only concerned with their differential coefficients, which are not altered by the addition of constants.) Thus our equation to the free surface now becomes

$$\eta = \frac{A - B}{V}\cos kx \quad \ldots\ldots\ldots\ldots\ldots\ldots\ (37)$$

so that

$$\eta_0 = \frac{A - B}{V}$$

(*b*) *The component of velocity normal to the bottom must vanish.*—This may be expressed either as $\partial\phi/\partial y = 0$ at the bottom, which we may suppose to be at $y = -h$, or as $\partial\psi/\partial x = 0$ at the bottom so that ψ is constant along the bottom, that is *the bottom is a streamline*. This condition gives

$$0 = \left(\frac{\partial\phi}{\partial y}\right)_{y=h} = \sin kx\,(Ae^{-kh} - Be^{kh}) \ldots\ldots\ldots\ (38)$$

or

$$A = Be^{2kh}.$$

(c) *The pressure just above the free surface must be constant and equal to the pressure of the atmosphere.*—The effect of surface tension (see Chapter IX), is to introduce a pressure difference $T\frac{\partial^2\eta}{\partial x^2}$ between the regions just above and just below the surface, the pressure being greater on the side towards which the surface is concave. The pressure just below the free surface is thus equal to (atmospheric pressure) $-T\frac{\partial^2\eta}{\partial x^2}$. Inserting expression (32) for ϕ in equation [23(a)] and remembering that $\partial\phi/\partial t = 0$ everywhere (since conditions are steady), we get, since the potential energy Ω is equal to the gravitational potential energy per unit volume $\rho g y$,

$$P_{atm} - T\frac{\partial^2\eta}{\partial x^2} = \text{Constant} - \frac{\rho}{2}\left[\left(\frac{\partial\phi}{\partial x}\right)^2 + \left(\frac{\partial\phi}{\partial y}\right)^2\right] - \rho g y$$

$$= \text{Constant} - \rho/2\,[\{V - k\,(Ae^{ky} + Be^{-ky})\cos kx\}^2 + \{k\,(Ae^{ky} - Be^{-ky})\sin kx\}^2] \quad \ldots\ldots (39)$$

At the free surface we replace e^{ky} and e^{-ky} by unity as before, and we also neglect terms of the order of k^2A^2 and k^2B^2 compared with VkA or VkB. Equation (37) shows that the ratio of such terms is of the order of magnitude $\eta_0 k$ which we have already assumed to be small. From equation (37) we also have

$$\frac{\partial^2\eta}{\partial x^2} = -\frac{k^2(A-B)}{V}\cos kx.$$

To this approximation, equation (39) can be satisfied if we equate coefficients of $\cos kx$ on both sides.

[Equating the terms independent of $\cos kx$ gives us no further result of interest, since equation (39) contains an additive constant, because of the presence of the arbitrary function of time $G(t)$ in equation (23).] Equating coefficients of $\cos kx$ gives us

$$\frac{Tk^2(A-B)}{V} = \rho Vk(A+B) - \rho g(A-B)$$

and using equation (38) to eliminate the ratio A/B, we get finally

$$V^2 = \left(\frac{g}{k} + \frac{Tk}{\rho}\right)\tanh kh \quad \ldots\ldots\ldots\ldots\ldots (40)$$

which is the general relation between V and k for which we are looking. To the approximation to which we are working ($\eta_0 k \ll 1$), we see that V only depends on the ratio of A to B, and not on A or B separately, that is, *V is, to this approximation, independent of η_0, the amplitude of the wave.* If we take terms of order $\eta_0 k$ into account this ceases to be true, and the velocity becomes a function of the amplitude, and *it is no longer possible for a sinusoidal wave to be propagated without change of form.* The same complications appear when we attempt to study sound-waves of finite amplitude (p 174).

For the present we shall consider the consequences of equation (40) in a number of special cases. It is interesting to have also the expressions for the particle velocities v_x, v_y as functions of the depth. These can be obtained by differentiating expression (32) with respect to x and y respectively, and using equations (37) and (38) to determine A and B in terms of η_0, V, k, and h.

$$v_x = -\frac{\partial\phi}{\partial x} = V - \frac{k\eta_0 V}{1 - e^{-2kh}}(e^{ky} + e^{-ky-2kh}) \cos kx,$$

$$v_y = -\frac{\partial\phi}{\partial y} = \frac{k\eta_0 V}{1 - e^{-2kh}}(e^{ky} - e^{-ky-2kh}) \sin kx \quad \ldots . \quad (41)$$

which expressions may be written also in terms of hyperbolic functions thus

$$v_x = V - \frac{k\eta_0 V \cosh [k(y+h)] \cos kx}{\sinh kh}$$

$$v_y = \frac{k\eta_0 V \sinh [k(y+h)] \sin kx}{\sinh kh} \quad \ldots\ldots\ldots . \quad (42)$$

with

$$V^2 = \left(\frac{g}{k} + \frac{Tk}{\rho}\right) \tanh kh \quad \ldots\ldots\ldots\ldots\ldots . \quad (43)$$

In interpreting these expressions it must be remembered that y is always *negative*, $y = 0$ corresponding to the free surface and $y = -h$ to the bottom, while V is the velocity we have superimposed upon everything to get steady conditions. The following are the most important special cases:

$kh \gg 1$. *Deep-water waves and ripples (depth large compared with wavelength).*—Here we may neglect e^{-kh} compared with unity, so that cosh kh and sinh kh are nearly equal, and replace tanh kh by unity. Equation (41) now shows that the amplitudes of horizontal and vertical oscillations are nearly equal, and 90° out of phase (the term V in v_x arising from the superimposed velocity), and that they both fall off with increasing depth according to the factor e^{ky}. That is, the paths of typical elements of fluid are *circles*. The velocity is given by

$$V^2 = \left(\frac{g}{k} + \frac{Tk}{\rho}\right)$$

which has a minimum value for $k^2 = \rho g/T$. Inserting for water the values $g = 981$ cm./sec.², $T \simeq 70$ dynes/cm., we get k min $\sim 3{\cdot}74$ corresponding to λ min $\simeq 1{\cdot}7$ cm. and V min $\simeq 27$ cm./sec.

Thus, for any liquid, there is a minimum velocity of propagation of waves, and for wavelengths below the critical value the velocity of propagation increases as the wavelength falls, that is, with increasing frequency, and such disturbances are usually called *ripples*. Since, in this region, V is effectively proportional to $\sqrt{T}$, we have here the basis of a very useful means of measuring surface tension. The measurement is usually carried out by disturbing the surface by a stylus vibrating at a known frequency ν, and measuring the *wavelength* of the resulting waves by photographic or optical methods, rather than by measuring V directly. ($\lambda\nu = V$ for any wave-motion.) It is, in practice, difficult to produce ripples of wavelength less than about a millimetre, not only because they are very rapidly damped out by viscosity, but also because the energy is carried away in the form of sound.

If λ is greater than the minimum value found above, then the velocity of propagation increases as λ increases, and the disturbance is usually called a *deep-water wave*. These exist over a very wide range of wavelengths, ranging from a few centimetres, produced by dropping a stone into water, to hundreds of metres, such as occur under conditions of "swell" at sea. The damping due to viscosity is usually extremely small, so such waves are propagated for very great distances, the main cause of the decrease in amplitude being the spreading of the waves outwards from their source. Since the velocity of propagation increases as λ increases, it follows that, *e.g.*, the effect of a storm in the ocean will first show itself at a distant point in the form of a *swell of long wavelength*. The disturbance of the surface due to the storm can be analysed into a frequency spectrum by means of Fourier's theorem (see Appendix), and it is the oscillations of longest wavelength and lowest frequency that travel fastest. It is not true that an oscillation of *zero* frequency would be propagated with infinite speed, because if the frequency becomes very small we must take account of the effect of the depth of the ocean in the manner to be explained below:

$kh \ll 1$. *Tidal waves. Depth small compared with wavelength.*—

$$\text{Since } \tanh kh = kh - \frac{(kh)^3}{3} + \frac{2}{15}(kh)^5 - \ldots \text{ for } kh < 1,$$

we may, without much error, replace $\tanh kh$ by kh as long as kh is appreciably less than unity. The expression (40) now becomes $V^2 \simeq gh$, the effect of the surface tension term being nearly always negligible. Thus, in this region of wavelengths, the velocity of propagation is no longer determined by the wavelength, but by the depth of water. The main application of the theory of this type of wave is to a discussion of the *tides* and to the multitude of subsidiary tidal phenomena that occur in straits and estuaries.

The tides may be regarded as a disturbance of the sea with an effective wavelength of the order of thousands of miles, which is very large compared with the depth of any ocean. As is well known, they originate in the gravitational attraction of the sea by the sun and moon, which, together with the rotation of the earth, produces a periodic rise and fall of sea-level at any given place in the sea. If the earth were entirely covered with water, the equilibrium configuration due to the attraction of the moon alone would be as shown in Fig. 61, the water being "heaped up" in the regions nearest to and furthest away from the moon. (It is, at first sight, difficult to understand why there should be this "heaping up" on the far side of the earth, as one would intuitively expect only a single "bulge", pointing towards the moon, and this is indeed what would happen if the earth and moon were held fixed. In fact, however, the earth and moon are being accelerated towards one another by their gravitational attraction, and this results in their both describing orbits about their common centre of gravity. Thus the *acceleration* produced by the moon's attraction is greatest for particles of water on the near side of the earth, slightly less for the earth itself, and least for particles of water on the far side, so that the latter tend to lag behind the earth.) Thus, if the earth were entirely covered with water, there would be a periodic rise and fall in its level at any given point as the earth rotated. The effect of the sun would be to increase the apparent amplitude when the sun, moon, and earth were nearly in line, and to diminish it when the sun and moon subtended an approximate right angle at the earth.

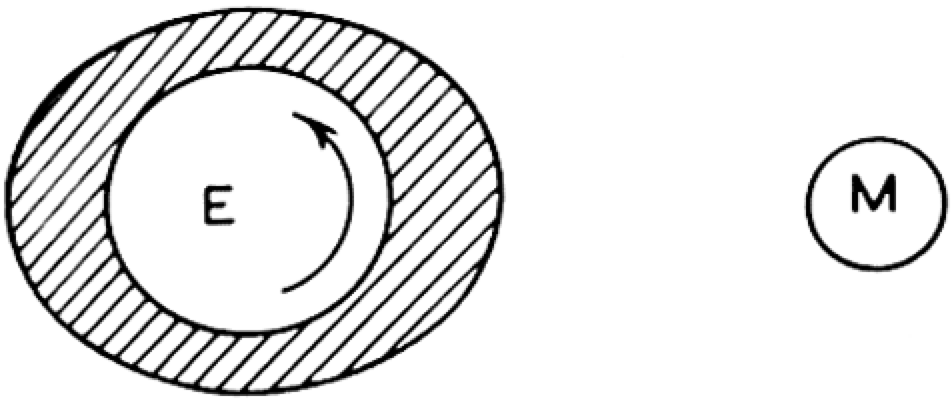


FIG. 61.

Although the earth is three-fourths covered with water, the true state of affairs is very much more complicated than that shown in Fig. 61, because the continents act as barriers, so that the ideal state of affairs shown in Fig. 61 can really only develop in the Southern Ocean, where there is a belt of sea all round the earth. Thus, we may think of the levels of the South Atlantic and South Pacific oceans being raised and lowered twice nearly every day, and these changes of level being propagated northwards. The periodic time is not exactly a day, because the moon is travelling round the earth and the tides consequently have a period of about $24 \times 28/27 = 24{\cdot}9$ hours. At any given place in the sea, the rate of propagation of the wave varies, as we have seen, as the square root of the depth. While, therefore, it is, in general, true that the phase of the tides lags more and more as one travels northwards, this phase is by no means the same at all places in the same latitude, both because the depth of the oceans themselves is not uniform, and also because of the individual peculiarities of the coastlines.

Of more interest than the *phases* of the tides are their relative *heights* in different places. These vary widely from about 1-2 ft in mid-Pacific to 40 ft in the Bay of Fundy. It is possible to understand this as another effect of the varying depths of the oceans and estuaries. If a wave is being propagated along a canal that becomes progressively shallower, the amplitude increases as the depth diminishes. Crudely, one may think of the waves in the shallow parts being overtaken, and thus reinforced by the faster waves from the deeper parts, although an exact treatment of even the simplified case of a bottom of uniform slope is very difficult. Another way in which one may come to the same qualitative conclusion is from a consideration of the energy transported by waves over water of different depths (Fig. 62). For a fixed *frequency* of wave, the velocity of propagation is less and therefore the wavelength must be shorter in the shallow region, and the potential energy associated with each wave is therefore less than in the deeper region. The rate of transport of energy cannot therefore be constant unless the amplitude is greater in the shallow region. This argument is not perfect, because it omits to take account of the fact that considerable *reflection* of any type of wave may well occur at a "step", but, by a well-established physical principle, this reflection becomes very much less if the "step" is replaced by a gradually sloping bottom.

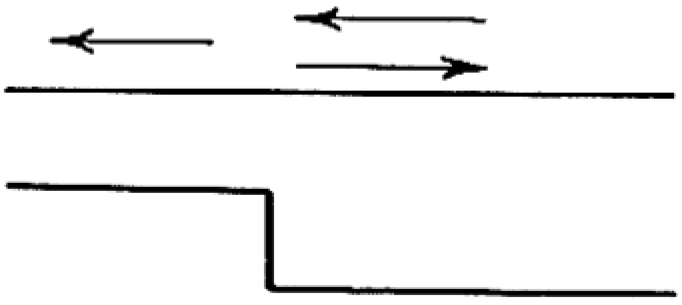


FIG. 62.

In this way it is possible to understand the amplification in the height of the tide as the tidal waves approach the land over the continental shelf, and further amplification may be produced by a rapidly narrowing channel such as the English Channel. The effect of depth on the rate of wave-propagation is also made manifest in the behaviour of waves on a sloping beach. It is a matter of common observation that the wave-crests are practically parallel to the shore, whatever may be the direction of the wind, and this can be understood as a *refraction* effect, the portion of the wave nearest the shore advancing the slowest on account of the smaller depth. As the wave approaches the shore, its amplitude increases for the reasons just mentioned and usually the wave "breaks" when its height becomes comparable with the depth of water, in which case the equations (32) and (33) no longer represent an acceptable solution of the hydrodynamic equations, because terms involving η_0^2 can no longer be neglected.

Conditions at the sea bottom.—Inspection of equation (42) shows a fundamental difference between tidal waves and deep-water waves. The amplitude of the vertical oscillations v_y is necessarily zero for both types of wave, but the ratio

$$\frac{\text{Horizontal amplitude at bottom}}{\text{Horizontal amplitude at surface}} = \frac{\cosh\,[k\,(y+h)]_{y=-h}}{\cosh\,[k\,(y+h)]_{y=o}} = \frac{1}{\cosh kh}.$$

For deep-water waves, kh is large compared with unity and this ratio is extremely small, the disturbance being appreciable only at distances below the surface comparable with the wavelength, but for tidal waves kh is small compared with unity, and the horizontal amplitude is practically independent of depth. Our solution thus implies *slipping of the water over the bottom,* whereas we know physically that the layer of water immediately in contact with the bottom is at rest, so that we must now ask whether this fact vitiates the account of tidal waves that we have just outlined. What seems to happen in practice is that vortices form in the boundary-layer near the bottom and act somewhat in the manner of *roller-bearings.* Since the motion of the water is oscillatory, the vortices probably also oscillate in position, but do not migrate continuously as they do in the wake behind an obstacle. The corrugations often seen on sandy beaches are probably induced in this way.

6. Non-linear hydrodynamics

Even if we assume our fluid to be Newtonian, or neglect viscosity altogether, there are still a great many problems in hydrodynamics that cannot be dealt with properly by the linear treatments that we have outlined. Neglect of the non-linear terms may have more serious results than simply spoiling the numerical correctness of a solution when the amplitudes become large, it may also lead to the overlooking of definite new effects.

There are two distinct reasons why non-linear effects appear in hydrodynamics, first the Bernoulli effect that we have already mentioned, arising from the fact that particle accelerations are not precisely equal to the partial derivatives $\partial v_x/\partial t$, etc., secondly the fact that the compressibility of all substances decreases with increasing pressure, so that the velocity of sound in an equation such as (19) p. 156, cannot be treated as a constant. The first consequence of these non-linearities is the so-called "overtaking effect", that is to say, the local velocity of sound is greatest where the pressure is highest, so that the high-pressure portions of a pulse or wave of large amplitude tend to travel faster than the low-pressure portions, and we get a progressive *steepening* of the profile of a pulse or wave-form, and it is no longer true that a sound is propagated without change in shape. Fig. 63 shows three successive stages in the propagation of a pulse of moderate amplitude, the leading edge becoming progressively steeper, while the low-pressure regions travel slowest and form a "tail". This change in profile can be followed theoretically by a method devised by **Riemann.** If we confine ourselves to motion in one dimension, equations (1) and (2) can be written, without approximation, as

$$v_x \frac{\partial \rho}{\partial x} + \rho \frac{\partial v_x}{\partial x} + \frac{\partial \rho}{\partial t} = 0 \quad \ldots\ldots\ldots\ldots\ldots \quad (44)$$

$$\frac{\partial v}{\partial t} + v_x \frac{\partial v_x}{\partial x} + \frac{1}{\rho}\frac{\partial P}{\partial x} = 0 \quad \ldots\ldots\ldots\ldots\ldots \quad (45)$$

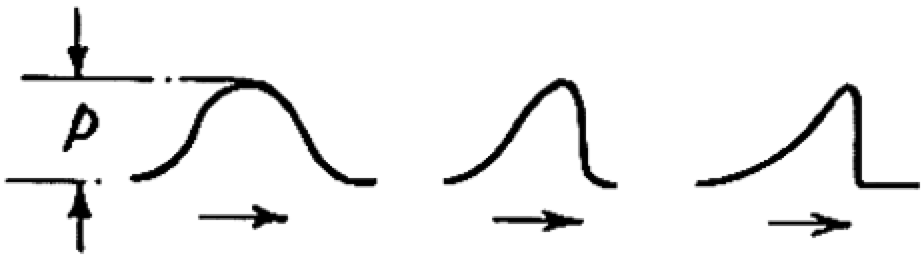


FIG. 63.

Since c, the velocity of sound, is given by $(\partial P/\partial \rho)^{\frac{1}{2}}$, we may write the last term as $\frac{c^2}{\rho}\frac{\partial \rho}{\partial x}$. These equations can be put into a more symmetrical form by introducing an auxiliary function, also depending on the pressure or density, known as the Riemann function. It is usually written σ, and is defined as follows:—

$$\frac{d\sigma}{d\rho} = \frac{c}{\rho} \quad \text{or} \quad \sigma = \int \frac{1}{\rho}\left(\frac{dP}{d\rho}\right)^{\frac{1}{2}} d\rho \quad \ldots\ldots\ldots\ldots \quad (46)$$

so that σ is, apart from an additive constant, a known function of ρ or P if the equation of state is known. We may then write

$$\frac{\partial \sigma}{\partial x} = \frac{c}{\rho}\frac{\partial \rho}{\partial x} \quad \text{and} \quad \frac{\partial \sigma}{\partial t} = \frac{c}{\rho}\frac{\partial \rho}{\partial t}.$$

The equations of motions (44), (45) become

$$\frac{\partial \sigma}{\partial t} + v_x \frac{\partial \sigma}{\partial x} + \frac{\partial v_x}{\partial x} = 0$$

$$\frac{\partial v_x}{\partial t} + v_x \frac{\partial v_x}{\partial x} + c\frac{\partial \sigma}{\partial x} = 0.$$

On adding and subtracting these two equations, we get

$$\frac{\partial}{\partial t}(\sigma + v_x) + (c + v_x)\frac{\partial}{\partial x}(\sigma + v_x) = 0 \ldots\ldots\ldots \quad (47)$$

$$\frac{\partial}{\partial t}(\sigma - v_x) - (c - v_x)\frac{\partial}{\partial x}(\sigma - v_x) = 0 \ldots\ldots\ldots \quad (48)$$

which are usually known as *Riemann's equations of motion.* They represent the generalisation to finite amplitudes of one-dimensional acoustic theory. Equation (47) states that the quantity $\sigma + v_x$ is being momentarily propagated with a velocity $c + v_x$, while the quantity $\sigma - v_x$ is being propagated with a velocity $c - v_x$. It is only possible to obtain analytic solutions of these differential equations in special cases, so that they have, in general, to be solved numerically, because the quantities c, σ, v_x, are varying from place to place, and from instant to instant. They do, however, show the overtaking effect very clearly. For, according to equation (47), the part of a pressure-pulse for which P, and therefore σ, is a maximum is being propagated with the greatest velocity, first because c is largest, secondly because of the presence of the v_x in the velocity of propagation $c + v_x$. In a sound pulse, v_x will usually be, as in equation (47), in the direction of travel of the sound. Equation (48) describes the opposite case, in which the particle velocity and direction of propagation are in opposite senses. These two types of wave or pulse frequently exist together, but it is no longer possible, as it is in ordinary acoustics, to superpose two such waves without their affecting one another.

In practice, the effect of the increase of c with pressure is usually much greater than the effect of the term v_x in $c + v_x$, so that the former is the main cause of the overtaking effect. The question now arises what happens when the leading edge of the pulse becomes vertical (Fig. 63), or nearly so. It is evident that it cannot become *quite* vertical, for this would imply an infinite rate of dissipation of energy by viscosity and thermal conductivity, however small these coefficients were numerically. In practice, it is found that, *e.g.*, an explosion pulse does have an extremely steep leading edge (whether propagated in air or in water), and, while it is probably correct to neglect dissipative effects in the remainder of the pulse, they cannot be neglected in the very steep region. This steep region is usually known as a *shock-front,* and it is often thought of as an abrupt jump of pressure.

A true jump is impossible on any molecular hypothesis and one ought to think of the change in pressure being spread over a distance of the order of 10-100 molecular distances. Numerical estimates of the effect of thermal conductivity and viscosity on the propagation of

Riemann waves do show that the dissipation of energy does become important for a shock-front of about this steepness.

It was shown by **Rankine** and **Hugoniot** that, if such a shock-front exists, the velocity of propagation, the stream velocity, and the height of the pressure-jump can be uniquely related to one another and that the rate of dissipation of energy during the propagation of the shock-front can be computed *without actual knowledge of the numerical values of viscosity and thermal conductivity*, it being assumed only that they are finite but small. This result is often found rather hard to believe, and the conclusion just mentioned was accepted only after a great deal of controversy.

A result showing a similar feature is that the energy dissipated when two electrically charged bodies are connected by a wire can be calculated even though the resistance of the wire is not known. Altering the resistance of the wire merely alters both the current and the time for which it flows in such a way that the energy dissipated is unaltered. Similarly, if it were possible to alter the viscosity of a liquid without changing its other properties, the effect would be to change the *slope* of the shock-front in such a way as to keep the dissipation of energy the same, without changing the shock-front equations themselves.

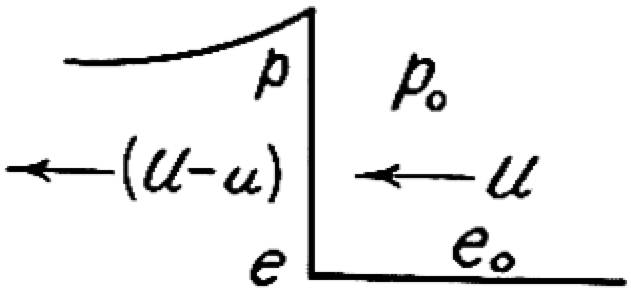


FIG. 64.

We now derive these equations. We suppose that the pressure in the undisturbed fluid is P_0, and that a shock is moving forwards into it with a velocity U, and that the stream-velocity immediately behind the shock is u. We let the densities in front of and behind the shock be ρ_0 and ρ respectively. It is convenient to impose a velocity U on everything, so that the shock is stationary, and the undisturbed fluid is arriving at the shock with the velocity U and leaving it with the velocity $U - u$ (Fig. 64). Then by conservation of matter we have

$$\rho\,(U-u) = \rho_0 U \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (49)$$

In unit time, an amount of momentum $\rho_0 U \times U$ is brought up to unit area of the shock, while an amount $\rho\,(U-u) \times (U-u)$ is carried away from it. The difference between these two quantities must, by the second law of motion, correspond to the net *force* acting across unit area of the shock, that is to $P - P_0$, so that

$$P + \rho\,(U-u)^2 = P_0 + \rho_0 U^2$$

or

$$P - P_0 = \rho_0 U u \quad \text{using (49)} \quad \ldots\ldots\ldots\ldots \quad (50)$$

A third equation can be obtained by an argument similar to that used in the discussion of the Joule-Thomson effect for the passage of gas through a porous plug (Roberts, *Heat and Thermodynamics*). One may think of the undisturbed fluid being fed into the stationary shock-front at a pressure P_0, being suddenly compressed, and then travelling away from the shock-front at a pressure P. Since the pressures are steady, *the history of each element of the fluid is the same.* The work done *on* unit mass entering the shock, by the pressure P_0 at one side of the shock, is $P_0 \times$ volume of unit mass $= P_0/\rho_0$, while the work done *by* the same unit mass on leaving the shock-front is P/ρ. The difference between these two quantities is partly accounted for by a difference in the kinetic energy of unit mass on the two sides of the shock, but the remainder can only represent a change in the internal or thermal energy. (As in the Joule-Thomson experiment, the corresponding change in temperature differs somewhat from that associated with a *reversible* change of pressure from P_0 to P.) Since no energy can enter or leave the fluid, the work done on unit mass must be equal to its total change of energy, kinetic and thermal, that is

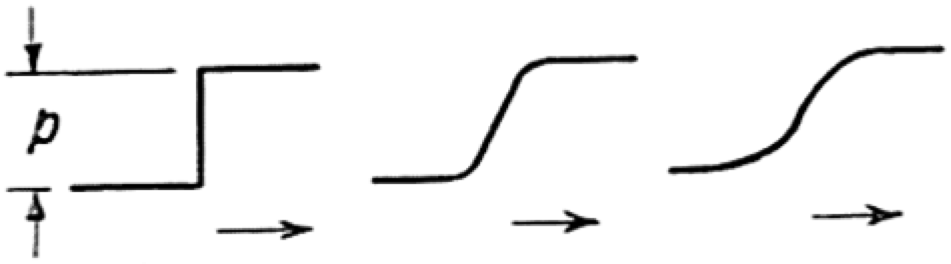


FIG. 65.

$$\frac{P_0}{\rho_0} + E_0 + \tfrac{1}{2}U^2 = \frac{P}{\rho} + E + \tfrac{1}{2}(U - u)^2 \quad \ldots\ldots\ldots \quad (51)$$

Providing we know the specific heats and the equation of state, which relates P, ρ, and T, equations (49)-(51) are sufficient to enable us to calculate P, ρ, U, and u, if *one* of these four quantities is known, together with P_0 or ρ_0 and T_0. These equations can be used to compute the rate of dissipation of energy associated with the shock, which, for weak shocks, is found to be proportional to the cube of the pressure-difference. It can also be shown that, for any ordinary substance in which $\partial P/\partial\rho$ increases with pressure, a "negative shock", $P < P_0$, is impossible, as it would imply the steady conversion of heat into mechanical energy, which is thermodynamically impossible. Moreover, it is obvious that, even if a sudden drop in pressure were produced artificially (Fig. 65), the regions of low pressure would be propagated *less rapidly* than the regions of high pressure, so that the discontinuity would tend to smooth itself off.

An interesting result can be deduced from equations (49) and (50), for weak shocks. We may write

$$u = \frac{U(\rho - \rho_0)}{\rho} = \frac{(U - u)(\rho - \rho_0)}{\rho_0}$$

from (49), so that

$$P - P_0 = \rho_0 U u = U(U - u)(\rho - \rho_0).$$

Now $(P - P_0)/(\rho - \rho_0)$ is, for moderate shocks, practically equal to $\partial P/\partial \rho$ or c^2, while, in the same circumstances u is small compared with U, and we have $U^2 - Uu = c^2$ or $U \simeq c + \frac{1}{2}u$, that is, the velocity of propagation of a shock-front is *intermediate between the local velocity of sound c, and the velocity c + u with which a Riemann wave of finite slope is propagated.* The shock-front travels slightly slower than a Riemann wave because shock-front theory takes the dissipation of energy into account, while the Riemann treatment neglects it.

Shock-fronts are produced by two main methods. First of all, *any* sound of finite amplitude will develop into a shock if it travels far enough, because the high-pressure regions will always overtake the low-pressure ones, but, in practice, this effect is significant only for the very intense sounds produced by explosions, where pressures are measured in kilograms or tons per square cm., rather than in dynes per square cm. as for ordinary audible sounds. Shock-fronts are also caused by anything that causes a fluid to move at a velocity greater than the local velocity of sound. This can be caused by a projectile, *e.g.* a bullet, or by the flow of the fluid itself, *e.g.* through a turbine jet. Usually, but not always, a shock-front divides a region of flow in which the stream-velocity is greater than the local velocity of sound from one in which it is less. Thus, in Fig. 64, the fluid is flowing *into* the shock at a velocity U, which as we have seen, is equal to $c + \frac{1}{2}u$, while it is *leaving* the shock at a velocity $U - u$, which is less than the local velocity of sound c. The phenomena associated with supersonic motion of projectiles and jets are extremely complex and will not be dealt with in this book.

Non-linear phenomena in surface waves. The "hydraulic jump".—We have seen above (p. 170) that sinusoidal surface waves can be propagated without change of form only if the second-order terms are neglected. If we take them into account, we reach the conclusion that *regions of elevation are propagated faster than regions of depression*, and it is natural to ask whether anything analogous to a shock-front exists. Plainly there is nothing corresponding to the increase of the velocity of propagation with pressure—the "overtaking effect" is here entirely due to the Bernoulli terms. Now it is a matter of observation that, in certain conditions, *e.g.* in a funnel-shaped estuary, an initially sinusoidal tidal wave can develop into a steep-fronted "wall" of water known as a "bore", and a theory quite analogous to that of shock-fronts can be developed. An analogue of a steady shock, resembling a small waterfall and known as the "hydraulic jump", sometimes occurs in rapid streams when there is a sudden change in the depth of water due, *e.g.*, to a submerged obstruction. The density of the water does not alter, but the stream velocities and depths of water on the two sides of the "jump" can be different, so that, in analogy with equation (49) we can write $u_1h_1 = u_2h_2$. Analogues of the other two equations can also be written down, but we shall not discuss this point further.

7. Conclusion

Hydrodynamics has the rather unjust reputation of being "untidy". While it is, as we have seen, usually true that problems that can be solved *exactly* are so idealised that they fail to represent any real situation, this is not unlike the picture presented by many other branches of theoretical physics. In such a problem as the behaviour of the tides, so many effects come into play, such as viscosity, the effect of the rotation of the earth, the varying depth of the ocean, etc., that any investigation is bound to be intricate if it is to give anything approaching a proper account of the actual situation. Again, experimental investigations of the situation in the neighbourhood of a projectile show that it is in fact extremely complicated, even for quite slow velocities of projection, and an exact solution of such a problem, even if it is ever found, would be almost useless as it would probably involve extremely complicated mathematical functions which would have to be tabulated, so that the present approach of seeking direct solution of the equations of motion by numerical methods is unlikely to be superseded in the near future. Moreover, the advent of electronic computing devices will certainly, within the next few years, enable vastly more complicated theoretical problems to be tackled efficiently and quickly, so that this gives a further advantage to the numerical approach.

BIBLIOGRAPHY FOR CHAPTER VI

H. Lamb, *Hydrodynamics*. 6th Edn. C.U.P. (1932).

L. M. Milne-Thomson, *Theoretical Hydrodynamics*. 2nd Edn. Macmillan, London (1949).

R. Courant and K. O. Friedrichs, *Supersonic Flow and Shock Waves*. Interscience Publishers, New York (1948).

E. G. Richardson, *Dynamics of Real Fluids*. Arnold, London (1950).

Examples on Chapter VI

1. Derive from first principles the Eulerian equations of hydrodynamics.

2. Derive an expression for the "hydrodynamic virtual mass" of a long cylinder, moving through a perfect fluid in a direction perpendicular to its axis.

(Hint.—Take coordinate axes at rest with respect to the cylinder. Assume a velocity potential $\phi = Ar \cos \theta + B \cos \theta/r$, where A and B are constants chosen to agree with the velocity at long distances from the cylinder and the boundary condition at its surface. Determine the corresponding velocities at each point in the fluid and transform to axes fixed in space. By integration over all space outside the cylinder the total kinetic energy per unit length of cylinder can be obtained.)

3. Describe the motion of a safety-razor blade as it falls through water, and give an explanation of the effects observed. Explain why the blade does not fall vertically "edgewise".

4. Draw a rough graph of the relationship between the force F acting on a small sphere and the velocity v with which it moves through an ordinary liquid or gas.

5. Show that the excess pressure p in a plane sound-wave of small amplitude obeys the approximate equation $\frac{\partial^2 p}{\partial x^2} = \frac{1}{c^2}\frac{\partial^2 p}{\partial t^2}$. In what ways does this equation fail to represent the facts if p is gradually increased?

6. Derive an expression for the velocity of propagation of waves on the surface of a sea of great depth.

7. A spherical gas-bubble is oscillating radially in a sea of infinite depth. Derive the equation governing the variation of the radius of the bubble with time. How is the flow round the bubble affected by the proximity of a free surface or a rigid bottom? (Use the method of images.)

CHAPTER VII

THE PHYSICAL PROPERTIES OF LIQUIDS

1. Introduction

We can describe a solid as an aggregate of molecules arranged, at random or in some definite lattice structure, so that the migration of a molecule to a new position in the solid occurs comparatively rarely. A gas can be thought of as a number of molecules distributed practically at random in empty space of volume very much greater than the total volume of the molecules. There is, however, no such clear-cut physical picture of a liquid. The total volume of a liquid is, as for a solid, not very much greater than the sum of the volumes of the molecules, but, on the other hand, the migration of molecules is much more rapid in liquids than in solids. Thus, we may attempt to describe a liquid as a solid whose lattice structure has been lost, or we may think of it as a highly compressed gas and attempt to devise a suitable generalisation of the kinetic theory of gases. Much effort has been spent on both these approaches, but no "model" is yet known that can be regarded as physically satisfying. On the other hand, several models are known that can each correlate a number of different properties of a liquid, and we shall describe a few of the most successful ones. We shall first summarise briefly the experimental facts that any ultimately successful theory will have to explain.

2. The observed properties of liquids

(*a*) *The distribution function.*—The first property that one thinks about is the way in which the molecules are arranged. They are not arranged in the completely random manner characteristic of a gas, nor are they arranged in a repeating lattice structure like a solid. If the scattering of X-rays by a liquid is determined, the pattern obtained is a series of *rings* rather than the isolated spots, characteristic of a large crystal. This pattern can be used to determine the structure of the liquid, the reader is referred, *e.g.*, to Frenkel's *Kinetic Theory of Liquids* for a description of the methods of interpretation, and the structure is usually expressed by means of the *distribution function*. This function is defined in the following way. We take a typical molecule, and plot, as a function of distance from its centre, the number of other molecules likely to be found within a small volume at various distances from our typical molecule. Plainly, if we suppose the molecules to be nearly rigid, the probability of finding the centre of a second molecule nearer to the centre of the first than the "gas-kinetic diameter" is negligibly small. On the other hand, if the molecules are closely packed, with very little "free volume", it is to be expected that there will be a great many pairs of molecules only a little further apart than one gas-kinetic diameter, so that we expect the distribution function to start at zero and rise steeply to a maximum. At great distances from our typical molecule, the probability of finding others within unit volume becomes practically uniform, and is proportional to the density of the liquid. The typical distribution function $g(r)$ inferred from X-ray data is plotted in Fig. 66. Theoretical investigations seem to show that something very like the observed distribution function is to be expected on purely geometrical considerations, a similar pattern being expected for rigid spheres packed at random.

(*b*) *Viscosity.*—The main characteristic of a liquid is that it cannot resist transverse shearing, but that a resisting force proportional to *rate of shear* is called into play. Gases also have viscosity, but this always increases as the temperature rises, whereas the viscosity of a liquid almost always falls rapidly with rising temperature. It has been customary for some time to attempt to analyse observed viscosities in terms of a law of the type $\eta \propto e^{U/kT}$, and, as we shall see, there is some theoretical basis for such a law, but it cannot be regarded as more than a first approximation and, even in allegedly "simple" liquids [the rather vague concept of an "associated" liquid will be discussed later (p. 265)], plots of $\log \eta$ against $1/T$ are often very far from straight lines. In some regions of temperature it is often found that a simpler law of the type $\eta \propto (1 + aT)^{-n}$, where a is a constant, gives at least as good a representation of the data as the exponential relationship.

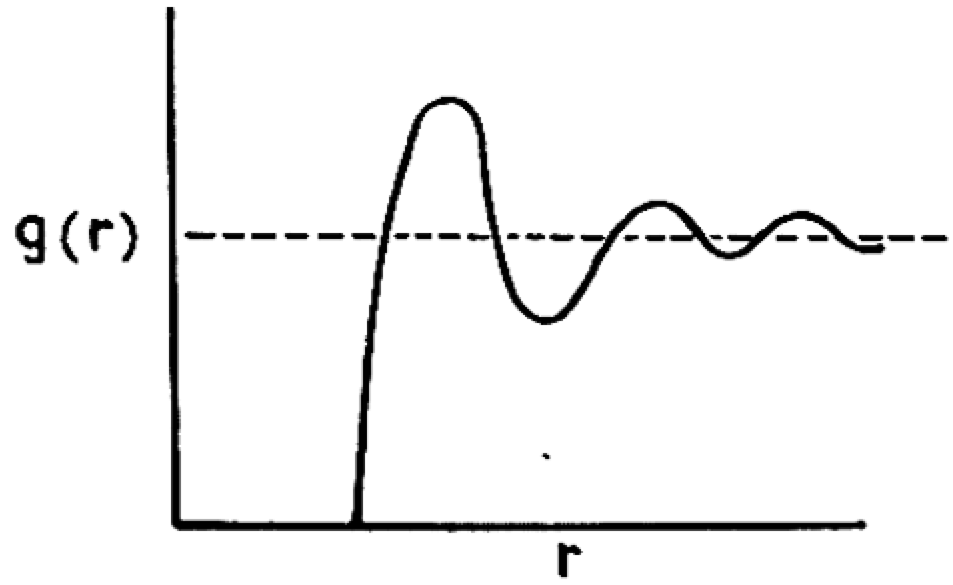


FIG. 66.

(*c*) *Surface tension.*—As is well known, a liquid behaves in many ways as if its surface were covered by an *elastic skin*, for example, a small drop of liquid is not pulled down into a flat cake by gravity but is nearly spherical. Again, it is possible with care to make a needle, or an aluminium disc, "float" on the surface of the water, although they are not capable of floating in the ordinary sense of the term and sink immediately the "skin" is broken through. Another familiar manifestation of surface tension is the continual motion of camphor chips on the surface of clean water. The surface tension of water is greatly lowered by dissolving camphor in it, so that a chip tends to be pulled *towards* places where the solution is weakest and the surface tension therefore strongest.

Such phenomena are usually interpreted by saying that a molecule at the surface of a liquid is subject to an asymmetric field of force, due to the attraction of the other molecules, because there is nothing balancing the attraction of the molecules *below* the one in question, so that there is a net force (Fig. 67) tending to draw a molecule from the surface back into the body of the liquid, where the net force on it will vanish by symmetry, and we may think that the free surface of a liquid has a certain *potential energy*, and that it is maintained in existence by thermal agitation, which brings more molecules to the surface as fast as others are dragged back into the main body of the liquid. There is no doubt that this view is correct, but some writers have criticised it on the grounds that, although it gives the required result that work has to be done to create a fresh surface, it cannot account for one of the common manifestations of surface tension, namely, forces acting *parallel* to the surface. According to Fig. 67, the force on a typical molecule can, by symmetry, act only *perpendicular* to the surface.

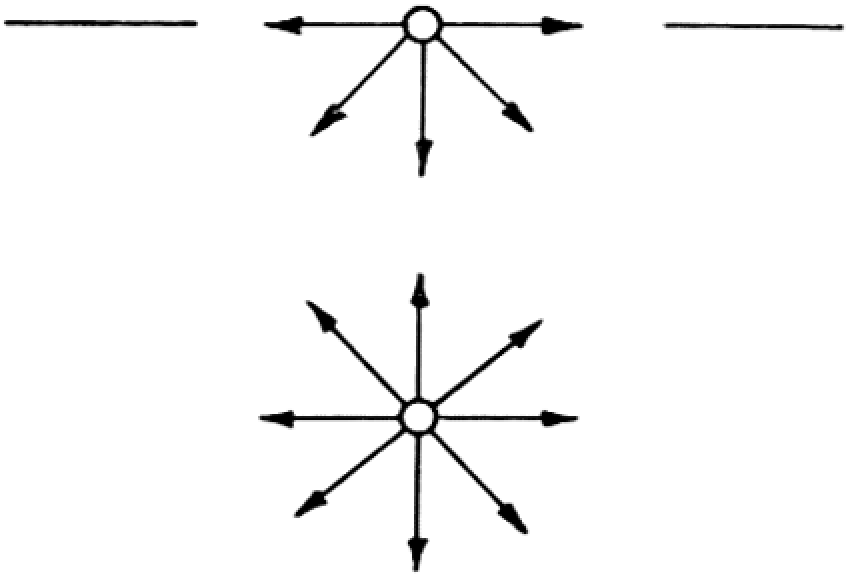


FIG. 67.

If we make a small framework with one movable side (Fig. 68), we may cover it with a soap film, or place it on the surface of a liquid, and then introduce another liquid into the region inside the frame, in which circumstances one observes "surface tension" forces acting parallel to the surface pulling or pushing the movable side (Fig. 68). Again, the rise of liquid in a capillary tube can be understood only on the basis that the surface tension forces have at least a component parallel to the liquid surface (Fig. 69). However, in Fig. 67, a symmetry exists between conditions on the left and right, whereas, to produce observable forces *parallel* to the surface, one must either introduce a solid boundary, or else a second liquid, and the symmetry between left and right is thus destroyed. In such cases,

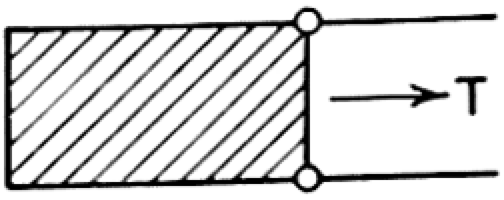


FIG. 68.

there is no longer any reason why the force on a typical molecule should be perpendicular to the surface.

The interactions of liquids with solids, and their spreading over one another, give rise to a very great variety of phenomena, some of which we shall meet later on (Chapter IX).

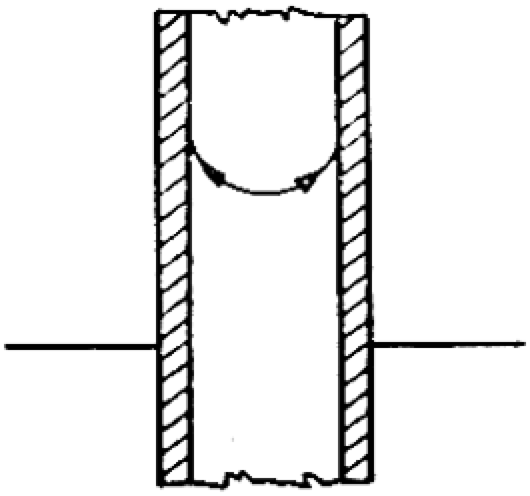


FIG. 69.

(*d*) *Tensile strength.*—It is not generally known that liquids can, under proper conditions, sustain very considerable tensions. There are various methods by which this can be demonstrated. One of the most direct is by spinning a J-tube containing the liquid about an axis perpendicular to the axis of the tube (Fig. 70), as was first done by **Reynolds.** The long arm may either be sealed under a vacuum, in which case the liquid is simply under its own vapour pressure, or else can be left open to the atmosphere. In either event, the rotation of the liquid about an axis implies an acceleration towards the centre of $\omega^2 r$ for each element of the liquid at a distance r from the axis, and this implies a *pressure-gradient* of $\rho\omega^2 r$ in the outward direction, so that the pressure decreases as one approaches the axis, and becomes negative there if ω is large enough. Usually tensions of the order of 5-10 atmospheres are enough to make the liquid rupture but **Briggs,** using special precautions, has shown that water and benzene will withstand tensions of the order of *hundreds* of atmospheres. Another way, due to **Berthelot,** in which tensions can readily be produced in a liquid is by sealing it up, leaving a small gas-bubble, in a tube which is then heated. The liquid expands faster than the glass and eventually dissolves the remaining gas-bubble and fills the whole tube. If it is now cooled the liquid is subjected to tension and eventually breaks with a "click" when the tube is cooled through about 5°-10° C. The corresponding tensions just before rupture are usually of the order of 50-100 atmospheres. Yet other ways in which a liquid can be subjected to tension are by high-frequency sound of high amplitude and by the Bernoulli effect associated with the rapid motion of liquids in pipes (p. 145) or with the motion of water in the neighbourhood of propellers. In these last three cases, the failure of the liquid and the subsequent

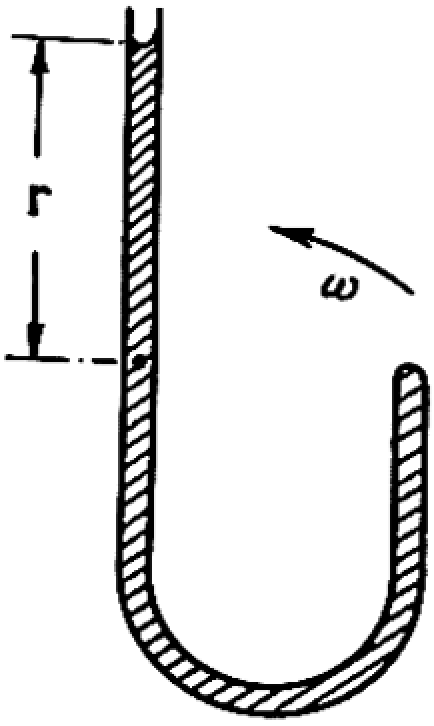


FIG. 70.

collapse of the vapour bubbles is known as *cavitation*, and is of considerable technical importance, as it not only limits the performance of propellers and of under-water signalling devices directly (one can say colloquially that cavitation causes a propeller to "lose its grip" on the water), but metals are rapidly eroded by cavitation, if it is at all intense or long continued, owing to the high pressures developed when vapour bubbles collapse to a very small volume.

(*e*) *The equation of state and specific heat.*—These properties are not peculiar to a liquid, being common to all matter, but any acceptable theory of a liquid must reproduce, *e.g.*, the compressibility of a liquid, at least in order of magnitude. It must also be compatible with the thermal measurements on the specific heat of the liquid and on its vaporisation and solidification, but discussion of these is outside the scope of this book.

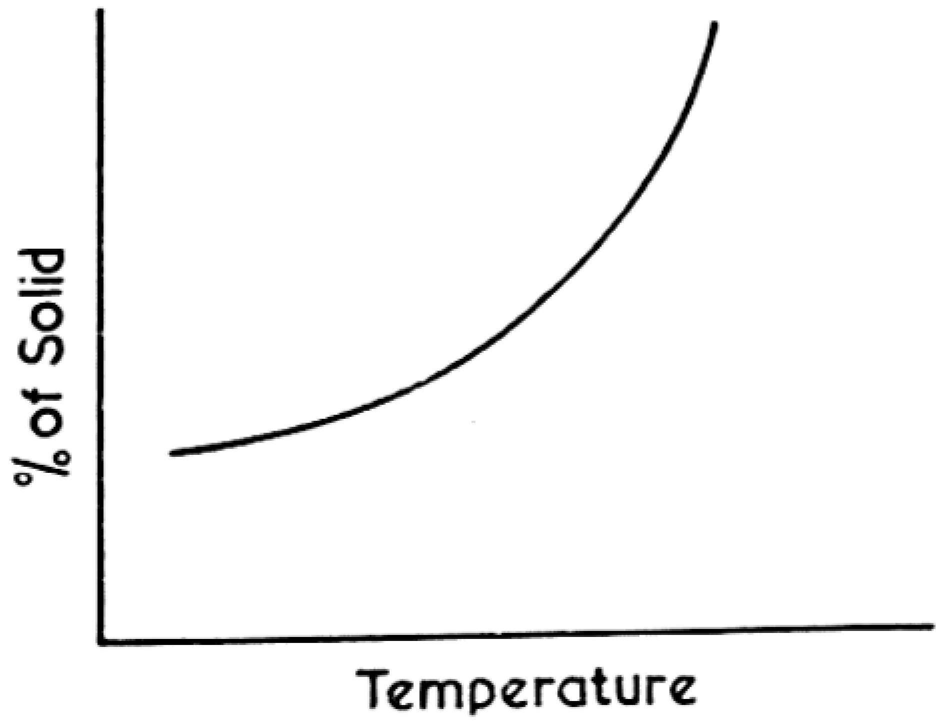


FIG. 71.

(*f*) *The "solution" of solids, liquids, and gases in liquids.*—It is an obvious fact that solids "dissolve" in liquids and solubilities in water range from minute fractions of 1 per cent., *e.g.* barium sulphate, to extremely high figures, *e.g.* sugar. [Indeed, crystals, *e.g.* of photographic "hypo" ($Na_2S_2O_3$), can be made into a "solution" without the addition of any water at all, for merely heating the crystals causes them to "dissolve" in the water of crystallisation already present in them.] Usually the solubility of a solid rises rapidly as the temperature increases, and the typical appearance of a solubility curve is as shown in Fig. 71 (*convex* to the temperature axis), though there are some solids, *e.g.* common salt (NaCl), for which the curve is nearly straight, and others, *e.g.* anhydrous sodium sulphate (Na_2SO_4), which become less soluble as the temperature rises.

An equally wide variety of behaviour is found for liquids. Some, *e.g.* mercury and water, are practically immiscible, others such as ethyl alcohol and water mix in all proportions, while yet others, *e.g.* ether and water, triethylamine and water, or phenol and water, show only limited mutual solubilities, and, for certain ranges of composition, two distinct layers are formed, one layer being a saturated solution of liquid A in liquid B, while the other layer is a saturated solution of liquid B in liquid A. If we define the solubility-curve as the relation between temperature and composition marking the limits of the two-layer regime, we usually get a curve of the type shown in Fig. 72 (mutual solubilities increasing with *rise* of temperature), *e.g.* phenol and water, but, here again, it is not unknown for the mutual solubilities to increase with *falling* temperature, *e.g.* triethylamine and water, or for the two-layer region to be bounded by a *closed curve*, *e.g.* nicotine and water.

All gases are to some extent soluble in liquids, indeed, it is doubtful whether a liquid sample entirely free from dissolved gases has ever been prepared. Apart from gases such as ammonia, which react chemically with the solvent, two fairly general statements about the solubilities of gases in liquids can be made. First, the solubility decreases as the temperature rises, secondly, the solubility expressed as mass of gas that dissolves in a given mass of liquid is, at a given temperature, nearly proportional to the pressure (**Henry's Law**). Since the volume of a given mass of gas is, by Boyle's law, inversely proportional to the pressure, Henry's law can also be expressed by saying that the *volume* of gas that dissolves in a given mass of liquid depends only on the temperature.

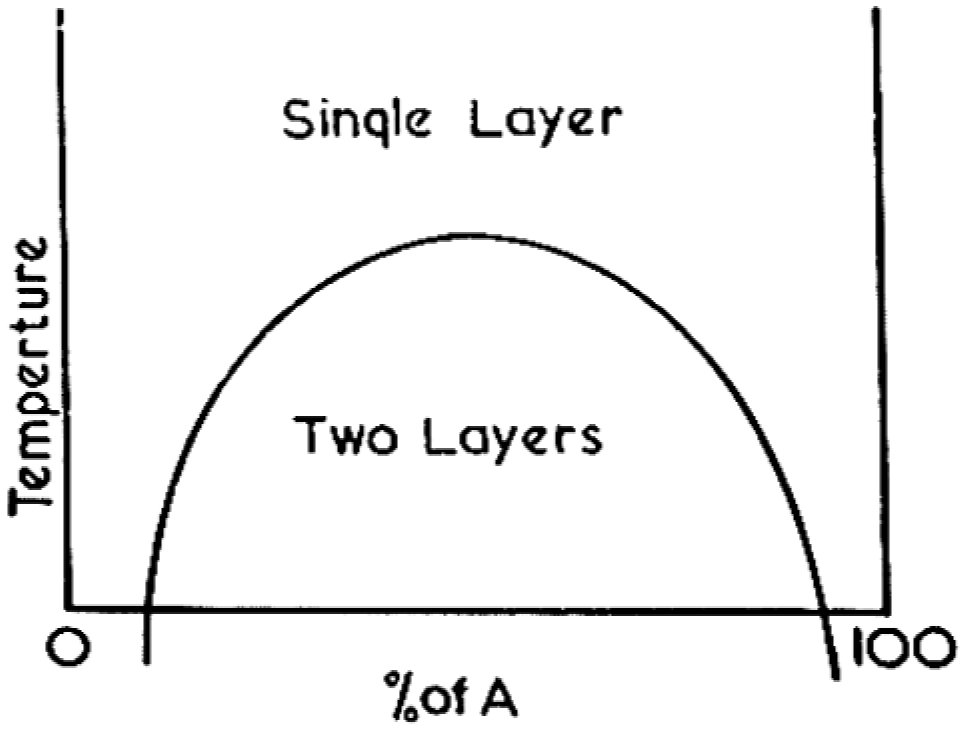


FIG. 72.

The properties of solutions will be discussed in more detail in Chapter VIII, but it is quite clear that both the existence of solutions, and the fact that the solubilities of chemically similar solids, *e.g.* barium and zinc sulphates, sometimes differ widely, are things that any ultimately successful theory of liquids will have to explain.

3. "Gas-like" models of a liquid

(*a*) *The van der Waals' model.*—The first successful attempt to describe a liquid was due to **van der Waals.** He began with the perfect gas model and then examined the effect of taking account of the molecular interactions. It has been shown that the form of the correcting terms is, to a first approximation, independent of the nature of the molecular interaction functions. We assume only that the force between two molecules is central (p. 8), and that it is repulsive at distances comparable with the gas-kinetic diameter, but becomes attractive at large distances, the potential energy curve being as shown in Fig. 73. It is to be expected theoretically that the attractive part of the energy (between molecules without dipole moments) should fall off as the *inverse sixth power* of the distance, the corresponding force sometimes being referred to as a van der Waals' force.

The physical reason for this can be understood if one thinks of the classical picture of an atom as a nucleus surrounded by one or more electrons revolving round it, which we may also think of as a *rotating electric dipole*. The field of such a dipole falls off as the inverse cube of the distance. Now suppose that we have a second atom in the neighbourhood. The second atom is polarised by the electric field of the first, and will now have a dipole moment which oscillates in phase with the rotations of the dipole associated with the first atom. If M_1 is the strength of the rotating dipole, the field at the second atom is proportional to M_1/r^3, where r is the distance apart of the two atoms, and the induced dipole moment is proportional to this field. Although the original dipole M_1 and the induced dipole, of strength proportional to M_1/r^3, are both zero when averaged over a long time (because of the revolution of the electron in a closed orbit), yet the time average of the *interaction* between these two dipoles is *not* zero, because the variations of the induced dipole moment are *in phase* with the inducing field, that is, with the original dipole moment. This interaction energy is proportional to *original moment* × *induced moment*$/r^3$, that is to r^{-6}, which is the physical origin of the inverse-sixth power law. There will, on the average, be no *direct* interaction between the two rotating dipoles because they will have slightly different periods of rotation, so the phase continually alters and the mean effect is zero. An *induced* dipole keeps a *constant* relationship in phase to the field that produces it. See Appendix, p. 288.

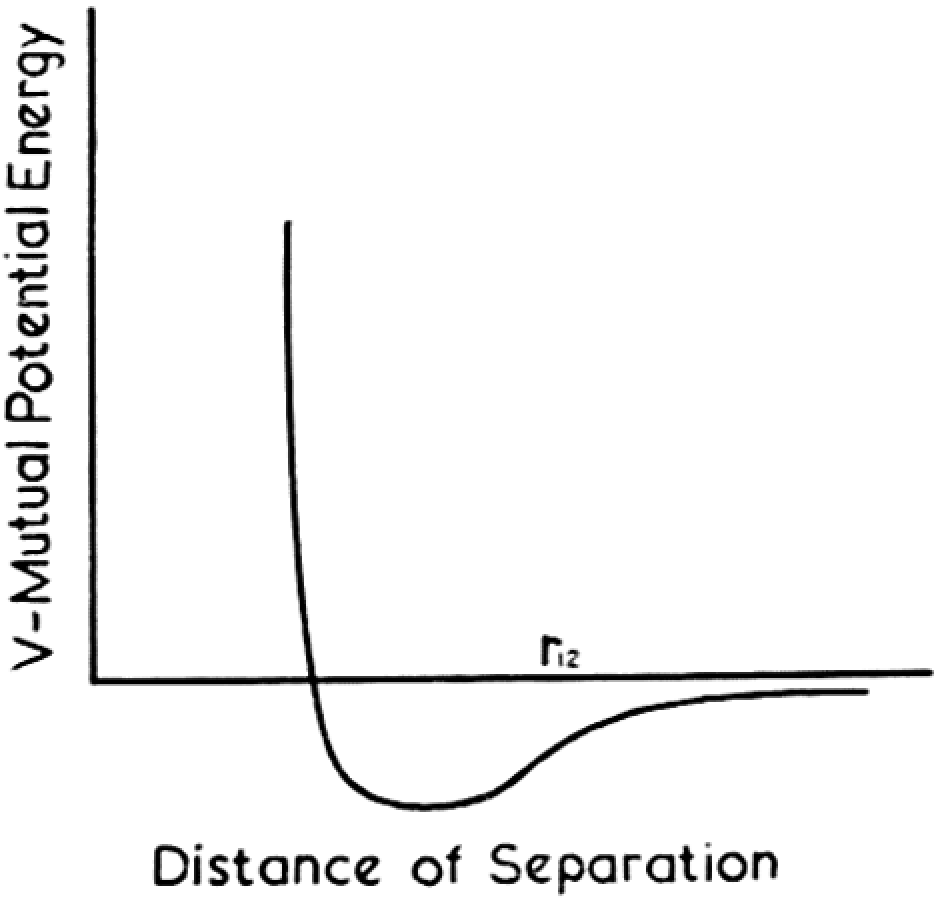


FIG. 73.

As for the repulsive forces, we must expect them to become important as soon as the outer electrons in two molecules approach one another. Now it is a fundamental feature of quantum mechanics that electrons are not to be thought of as localised in a very small volume, but as likely to be found with varying probabilities in various regions of space. Alternatively, we may think of an electron as represented by a certain distribution of electric charge permeating all space, but with density falling off very rapidly with distance. For an electron bound to a nucleus the corresponding probability or charge-density falls off *exponentially* with increasing distance. Thus, if we push two atoms together, the "charge-clouds" associated with the outermost electrons "overlap" more and more, and the repulsive forces between atoms result from the "deformation of the outer electronic orbits" or, looked at another way, from the electrostatic repulsion between the charge-clouds. This latter effect is roughly proportional to the density of the charge-clouds in the region of overlap, that is to say, to e^{-r/r_0}, where r_0 is of the same order of magnitude as the radius of the electronic orbits. Thus we arrive at the following approximate theoretical expression for the mutual potential energy of two atoms,

$$V = -\frac{M}{r^6} + Ne^{-r/r_0},$$

that is, we have an attraction at long ranges but a repulsion at short ranges. This is the simplest type of interaction that can possibly represent physical reality.

The rigorous evaluation of the equation of state of an assembly of molecules with such an interaction function is still an unsolved problem. What *can* be done rigorously is to evaluate the *second virial coefficient* as a function of temperature, that is the coefficient B in the equation of state expressed in the following form, which is the customary way of exhibiting data on the compressibility of gases:

$$Pv = RT + \frac{B}{v} + \frac{C}{v^2} + \dots$$

The theoretical prediction of C, or of any of the higher virial coefficients, is a matter of considerable difficulty, for it becomes necessary to consider in detail cases where two, three, four, or more molecules are in close proximity. In the kinetic theory of gases it is nearly always sufficient to confine oneself to binary collisions, but higher order collisions become more and more important as the gas is compressed, until the liquid state appears, when their effect becomes dominant.

We shall not attempt the rigorous treatment, but shall give a rough derivation of the form of the corrections to the perfect gas laws necessitated by interactions, which, although it is by no means free objection, does at least enable one to see the physical origin of the main correcting terms. We first attempt to take account of the repulsive forces, which we represent roughly by assigning a finite diameter to the molecules, comparable with the gas-kinetic diameter σ, and suppose that two molecules can never approach closer than this. Then we have to take account of the fact that the centre of a typical molecule moving in an enclosure of volume v cannot approach the centre of any other molecule closer than a distance σ. If we first suppose that all the molecules, except the typical one we are considering, are at rest, we should have to draw a sphere of radius σ round each of the other molecules, the total volume not accessible to our typical molecule, the "excluded volume", thus being eight times the total volume of the molecules or $\frac{4\pi N}{3}\sigma^3$. In the kinetic theory derivation of the equation $Pv = \frac{1}{3}M\bar{c}^2$ (p. 115) we neglected the finite volume of the molecules, and a correction must now be made for this by writing $v - f$ instead of v, where f is the excluded volume.

The figure of eight times the actual volume of the molecules that we have just deduced for f must now be modified to allow for the fact that a typical molecule does not merely move round at high speed, "bumping into" other nearly stationary molecules, in which case the above figure would be correct. In fact the "target" molecules are themselves moving at the same mean speed as the typical one. This necessitates a reduction of the excluded volume by a factor of two, a type of correction that we also met when discussing the relation between mean free path and molecular diameter (p. 123). For those "target" molecules nearly in contact with the walls, which are just in the process of rebounding, this modification is not applicable, instead, such a molecule contributes, on the average, only a hemisphere of radius σ to the excluded volume, not a whole sphere. We thus deduce that the excluded volume should be $\frac{2}{3}N\pi\sigma^3$, that is *four* times the total volume of the molecules.

We have also to correct for any possible attractive forces between the molecules. On the whole these will balance out, but, near the walls there is no longer a symmetrical distribution of molecules round a typical one. The momentum component, perpendicular to the wall, of an atom rebounding from it is thus reversed not only by the forces between wall and molecule, the force on the wall giving rise to the observed "pressure", but partly also by the intermolecular forces (cf. Fig. 67). Thus, this correcting term in the pressure is proportional first to the number of molecules which are striking the walls in unit time, that is, to the total number of molecules in the vessel, secondly, to the number of molecules near enough to a typical molecule to exert an appreciable force on it, which is also proportional to the density or number of molecules present. Thus, to a first approximation, the correcting term in the pressure is proportional to the square of the density, that is to $1/v^2$, and we correct our equation of state for the perfect gas by writing

$$\left(P+\frac{a}{v^2}\right)(v-b)=RT \qquad \ldots\ldots\ldots\ldots\ldots\ldots (1)$$

We shall now determine the consequences of this equation, without making any further attempt to improve the correcting terms. Theoretically, it is known that, even for a very simple type of molecular interaction function, the true correcting terms would become very numerous long before the liquid state was approached, so that the assumption that a and b can be treated as constant cannot be quantitatively correct. The arguments used in deriving equation (1) are crude, and can only be relied upon as first approximations.

Nevertheless, we shall now see that equation (1) does give a satisfactory qualitative account, not only of the phenomenon of condensation, but of some of the properties of the liquid state as well. One sometimes speaks of a *van der Waals' gas* (*or liquid*) by which we mean a hypothetical assembly whose molecular interactions are precisely described by equation (1). We develop some properties of this assembly and then compare them with those of actual gases and liquids.

Using equation (1) it is a simple matter to plot isothermals, for we may write

$$P=\frac{RT}{v-b}-\frac{a}{v^2},$$

and the form of this relation is sketched in Fig. 74 for a number of different temperatures. These are divided into two sets by the isothermal ACB, which has a point of inflection at C, the isothermals below it showing maxima and minima, while those above it do not. We may at once conclude, on physical grounds, that the rising portions, such as FG and LM of the isothermals, which would correspond to the volume increasing with increasing pressure, are unstable and cannot represent physical reality. It therefore seems that the assembly, on being taken along the path DEF, must, at some stage, jump discontinuously to the portion of the curve GHK, as the assembly has certainly become unstable by the time it reaches the point F. We are thus strongly tempted to identify these two portions of the curve with the isothermals of the liquid and vapour phases. We may then suppose that, in Fig. 74, the two phases are equally stable, at a given temperature, at some definite pressure corresponding to the line *EH*. An application of statistical mechanics, beyond the scope of this book, confirms this supposition, and indicates that the line *EH* must be drawn so that the area between *EH* and the volume axis is the same as the area enclosed between the curve *EFGH* and the volume axis. This result is known as the *rule of equal areas*. (Various attempts have been made to establish it without recourse to statistical mechanics, by thermodynamic reasoning alone, but the author believes them all to be fallacious.)

The portion of the line *EF* corresponds to a state of the assembly stable with respect to small fluctuations of pressure, etc., but less stable than a state in the region *HK*, to which it will go over spontaneously if a sufficiently large fluctuation occurs, or *if the assembly is "seeded" by the introduction of a suitable nucleus*. We are thus led to associate the portion *EF* of the isothermal with the familiar *superheated state* of a liquid, which occurs, *e.g.*, when one attempts to boil a liquid in a smooth glass vessel, which results in "bumping" or the sudden formation, with almost explosive violence, of large bubbles of vapour. This is obviated by introducing pieces of unglazed porcelain, which continually release small bubbles of gas capable of acting as nuclei and assisting the formation of vapour bubbles. The reason why it is difficult to create a small bubble of vapour in a liquid phase will become apparent later on when we consider surface tension (p. 245).

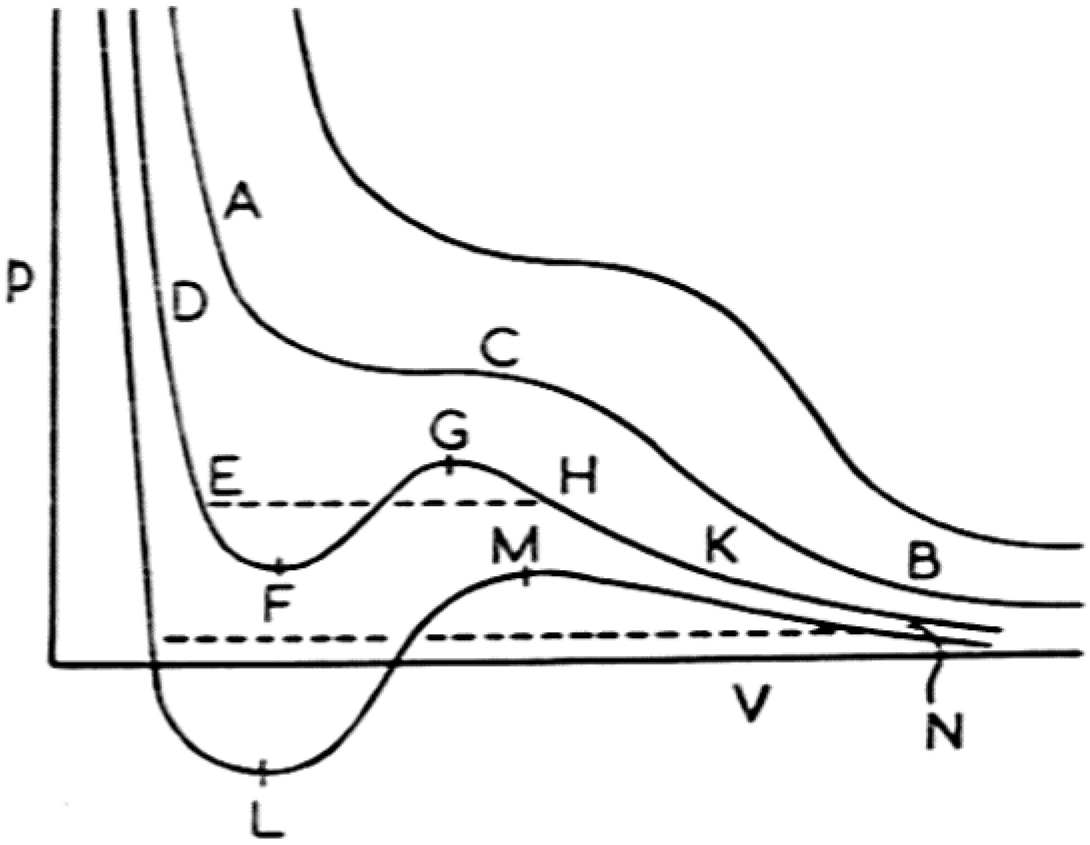


FIG. 74.

The portion *HG* of the isothermal corresponding to the vapour phase is to be thought of as a region where the vapour phase is stable against small fluctuations, but nevertheless, less stable than the liquid phase. Again, we know experimentally that such a situation can be realised. The *cloud-chamber,* due to **C. T. R. Wilson,** exploits the existence of this metastable region. Basically, it consists of a chamber in which the air is practically saturated with water vapour. By suddenly withdrawing a piston, the air can be expanded adiabatically to a volume that bears a predetermined ratio to its original volume, which results in a fall in temperature that more than compensates for the effect of fall in pressure, and matters are arranged so that the water vapour ends up in a state represented by a point on a portion such as HG of the isothermal. In this state the condensation occurs only if some kind of *nucleus* is present, to assist in the formation of an initial drop of liquid. Among bodies that are capable of acting as nuclei in this way are *dust particles* or *ionised gas molecules.* The action of an electrically charged particle is to attract neutral molecules of water vapour to it and thus to facilitate the formation of a drop of water, which the effect of surface tension tends to prevent. The electric charge on an ion, although small, nevertheless produces in its neighbourhood fields that are comparable with any electric fields that can be produced in the laboratory, and its effect in "gathering together" water molecules is very appreciable. By carrying out a few preliminary expansions and allowing the "fog" to settle, it is possible to remove all the dust particles originally present in the chamber, which is then ready to detect and make visible any charged ions that may be produced during the expansion, *e.g.* by the passage through the vapour of any high-speed charged particle such as an α-particle. Each ion produced by the passage of the α-particle produces a small drop of water, and thus renders the track of the α-particle visible. If the expansion ratio is too great, the chamber fills with fog, whether charged ions are present or not, a situation than can be described either by saying that we have gone past the metastable regions such as HG and have reached a state of affairs in which the vapour phase is definitely unstable, or else one may say that even uncharged molecules are capable of acting as "nuclei" if the degree of supersaturation is too great.

The fact that one has to *expand,* not compress the vapour in order to reach a saturated or super-saturated state is worth noticing. If a vapour is expanded *isothermally* it is always taken further from the liquid region, but, with some vapours, of which water is one, the fall in temperature during an *adiabatic* expansion more than compensates for the effect of the fall in pressure. This is sometimes expressed loosely by saying that "the specific heat of saturated steam is negative".

Thus, our simple model is capable of representing two of the properties of liquid and saturated vapour. By the help of the rule of equal areas we obtain a law of the variation of vapour pressure with temperature which is qualitatively the same as that for actual liquids. Moreover, a very simple argument enables us to locate the point C, the corresponding temperature being that above which no separation into two phases is possible. The coalescence of the maxima and minima occurring at lower temperatures implies the existence of a *horizontal tangent* and a point of inflection at C, so that we have

$$\left(\frac{\partial P}{\partial v}\right)_c = \left(\frac{\partial^2 P}{\partial v^2}\right)_c = 0,$$

which, using equation (2), give us easily

$$v_c = 3b, \quad P_c = \frac{a}{27b^2}, \quad T_c = \frac{8a}{27\boldsymbol{R}b}.$$

Thus, from the observed critical data, we can also calculate values of the constants a and b. We do not find more than qualitative agreement with values of a and b deduced from the virial coefficients, *i.e.* from the observed departure of the gas from Boyle's law at high pressures—and indeed, quantitative agreement is not to be expected, as the equation itself cannot theoretically be expected to hold in the neighbourhood of the critical point. In particular, the theoretical relationship between the three critical coefficients obtained by eliminating a and b,

$$\frac{P_c v_c}{\boldsymbol{R}T_c} = \frac{3}{8} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (2)$$

does not hold accurately, the observed ratios being usually in the neighbourhood of 0·2 — 0·3.

Yet another qualitative success of the van der Waals' equation is the following: below a certain temperature, the isothermals take a form resembling that of LMN in Fig. 74. The rule of equal areas still gives a positive vapour pressure for such an isothermal, but part of the metastable region now corresponds to the liquid being at a negative hydrostatic pressure, that is, *under tension*. We have already mentioned (p. 185) that liquids can, under proper conditions, be subjected to tensions of the order of hundreds of atmospheres, but reference to the original papers must be made for a detailed discussion of the various methods of applying and measuring these tensions. (It seems quite likely that the observed values correspond to the liquid "tearing away" from the walls of its containing vessel rather than "breaking" in the strict sense.) If one assumes reasonable values of a and b the theoretical tensile strength of a van der Waals' liquid at room temperature, *i.e.* the distance of the point L (marking the limit of the metastable region) below the volume axis is of the order of 500 atmospheres. This agrees with observation far better than the customary assumption made in the literature that the tensile strength is numerically the same as the so called "intrinsic pressure" a/V^2, which is usually of the order of 10,000 atmospheres.

Thus the value of the van der Waals' equation of state, and of other similar types of equation, is that they give a qualitative account of quite a number of phenomena, and at the same time enable us to make "order of magnitude" estimates of the van der Waals' constants a and b. The constant b is, as we have seen, a measure of the volume of a molecule, and estimates obtained from the departure of the gas from Boyle's law and from critical data can be compared with other estimates of atomic volumes, obtained from the measurements of thermal conductivity and viscosity of gases (p. 124), from X-ray scattering, which measures effectively the radii of the outer electronic orbits, or from quantum theoretical considerations, which also determine the effective radii of electronic orbits for fairly simple atoms and molecules. In almost every case one gets satisfactory "order of magnitude" agreement from all of these methods. The "gas-kinetic" and "van der Waals'" values for the effective radii are usually a little larger than the radii of the outer orbits. This is to be expected as any appreciable intermingling of the outer orbits of two molecules would bring into action large electrostatic repulsions. See Table VI (p. 315).

As we shall see in considering surface tension, various independent estimates of the van der Waals' constant a are possible, *e.g.* from the surface tension, or from the latent heat, which can be compared with those deduced from the departure from Boyle's law and from critical data. **Lennard-Jones** and his collaborators have made considerable use of a known result of statistical mechanics, first proved by **Ursell**, and quoted on p. 189, that the second virial coefficient and its variation with temperature can be calculated unambiguously if the interaction energy be supposed known as a function of the distance between two molecules, it being assumed also that the interaction forces are *central* and that the mere proximity of a third molecule does not appreciably modify the mutual potential energy of any two. For simple molecules, *e.g.* monatomic molecules of the rare gases, the interaction functions can be computed quantum-mechanically, and the second virial coefficients predicted from them, with fair agreement with experiment. This type of calculation thus supersedes older approximate attempts to estimate the van der Waals' constant a. Accordingly, one can accept with some confidence the information about the interactions of more complicated molecules obtained by a study of the second virial coefficient. The reader is referred to Fowler and Guggenheim, *Statistical Thermodynamics*, for details.

(*b*) *The "cell" model.*—The van der Waals' model cannot be expected to provide any information about such phenomena as viscosity and diffusion. Since we know from other evidence that multiple collisions play a large part in determining the properties of a liquid, it cannot be expected that any extension of the kinetic theory of gases, in which only binary collisions are normally considered, is likely to be fruitful. Indeed, we know (Chapter V) that the viscosity of a gas increases with rising temperature, whereas that of a liquid rapidly falls.

The "cell" or "cage" model of a liquid attempts to take account of the interactions by postulating that a typical molecule is confined in a "cage" of a volume that is comparable with its own, as a result of being surrounded fairly closely by other molecules Such processes as liquid flow, diffusion, and heat conductivity are to be thought of as a progressive transfer of the molecules from cell to cell in response respectively to space gradients of pressure, concentration, and temperature. One may suppose neighbouring cells separated by "potential barriers" of average height U, representing crudely the effect of the interactions, and an average molecule may be expected to make quite a number of collisions with others that form the walls of the "cell" before it acquires enough kinetic energy to escape. In equilibrium conditions the number of molecules which have a total energy of U or more, is by Boltzmann's expression, proportional to $e^{-U/kT}$, so that the number of molecules which leave their cells in unit time is proportional to $e^{-U/kT}$. Hence we might, at first sight, expect to find a factor of this kind in the coefficients of diffusion and thermal conductivity, and a factor $e^{+U/kT}$ in the coefficient of viscosity. In practice it is possible to express roughly the temperature variation of the coefficient of viscosity by means of such a function, but the coefficient of thermal conductivity does not seem to contain such a factor, while data for the variation with temperature of the diffusion coefficient in liquids are scarce. In any case, it is not to be expected that the three coefficients should vary in exactly the same way with temperature. The effect of a pressure-gradient may be thought of as that of *raising* the potential barrier on one side of the cell and *lowering* it on the other side. The effect of a concentration gradient is that the probable *number of molecules within a cell* varies with the distance, while the effect of a temperature gradient is that the factor $e^{-U/kT}$ and, therefore, the mean number of molecules *escaping from a cell*, is a function of distance. Various workers have derived quantitative formulae based on this model.

The quantity U, like the van der Waals' constant a, can also be related to the latent heat of evaporation and surface tension of a liquid, but this connection between U and a will be better understood when we consider the "hole" model of a liquid in the next article.

4. "Solid-like" models of a liquid

(*a*) *The "hole" model.*—Since the density of a liquid never differs by more than a few per cent. from that of its solid, in contrast with the enormously smaller density of the vapour, many workers have attempted to derive a theory of liquids by some modification of that of solids. One of the most promising of such developments is the "hole" theory. According to this theory, the melting of a crystalline solid is thought of as the progressive transfer, as the temperature rises, of molecules from their proper lattice sites to other positions in the lattice. When enough molecules have migrated to such interstitial sites, it is to be expected that the whole lattice structure will become unstable and break down, and a liquid is thought of as a sort of random mixture of molecules and unoccupied lattice sites or "holes". The "flow" of a liquid is then to be interpreted as a progressive process of molecules filling up neighbouring holes. To remove a molecule from the liquid, leaving a "hole", a quantity of work U must be done against the inter-molecular forces, and the number of "holes", and thus the rate of flow of the liquid, is thus again proportional to a factor $e^{-U/kT}$. So that, in practice, the predictions of the "hole" model are very similar to those of the cell model. However, the relationships between U, the energy necessary to create a "hole" by removing a molecule from the liquid, and the surface tension and latent heat, are much more obvious physically than when U is to be thought of as the height of the potential barrier that occurs in the "cell" model. This simple conception has led, in the hands of **Eyring** and **Lennard-Jones**, and others, to the interpretation of many of the properties of liquids, and to a development of a semi-quantitative theory of fusion on the lines sketched above. We may mention one most ingenious development of the theory by **Fürth**. He supposes that the energy necessary to create a "hole" may be represented adequately by the surface tension multiplied by the surface area of a "hole", while, as it moves through the liquid it may be supposed to have kinetic energy because of the hydrodynamic virtual mass (p. 164) associated with the fact that some liquid is displaced by the "hole". These ideas are almost certainly correct in principle, but the quantitative applications of them proved disappointing. This is probably because the conceptions of surface tension and virtual mass are essentially *macroscopic* ones applicable to a continuous fluid. As one might expect intuitively, the most probable size of "hole" turns out to be comparable with the volume of a molecule, so that strict quantitative application of these conceptions is not possible.

(*b*) *The Debye theory.*—Debye's very fruitful method of describing the properties of solids involved representing the sum total of the motion of individual molecules under thermal agitation by means of a superposition of sound-waves. Both compressional waves and shear waves (p. 107) were used, the possible frequencies of which depended on the acoustic modes of vibration of the solid, treated as a continuum. The amplitude of each mode was fixed by Planck's expression $\frac{h\nu}{e^{h\nu/kT}-1}$ for the mean energy of a harmonic oscillator of frequency ν at temperature T. In this way it is possible to give a very acceptable description of the thermal properties of a solid, and, by development of the theory, of its mechanical properties. Attempts have been made to apply the theory to liquids with some success, one of the main difficulties being to decide whether or not a liquid should be supposed capable of supporting shear waves (p. 108). Experimentally, ordinary liquids cannot support shear waves of low frequencies, but the frequencies that are most important in Debye's theory are comparable with the number of collisions suffered by a given molecule per second, that is of the order of 10^{12}/sec., which is far above any ultrasonic frequency that has been investigated experimentally. Attempts have been made to settle this question by analysis of the scattering of light by liquids, but the results are inconclusive. For a further discussion of this point the reader is referred to Frenkel's *Kinetic Theory of Liquids*. The propagation of shear waves in highly viscous *non-Newtonian* (see p. 140) liquids has been experimentally demonstrated, but no such result has yet been obtained for any Newtonian liquid.

5. Measurement of the physical properties of liquids

(*a*) *Elastic properties*.—The only elastic modulus that can be defined for a liquid is the bulk modulus K, defined in the same way as for a solid as $v\,(\partial P/\partial v)$. Early attempts to measure this quantity were not satisfactory, because the liquid has to be contained in some kind of vessel, and the expansion of the vessel when the pressure inside it is increased is comparable with the change of volume of the liquid. Even if the same pressure is applied inside and outside the vessel, the volume change of the vessel is *not* zero, but is now a decrease, the same as that which would occur if a piece of the solid that would exactly fit the inside of the vessel were subjected to the same pressure externally. The volume change to be expected can be computed theoretically for reasonably simple shapes of vessel, such as spheres, or cylinders with flat or hemispherical ends, but if glass vessels are used an additional complication arises because glass is not elastically isotropic. Metal vessels are more likely to be isotropic, but often show porosity.

Probably the first accurate determinations are due to **Regnault**, who surrounded the actual vessel containing the liquid experimented on by another vessel, so that the same pressure could be applied to the inside only, the outside only or to both the inside and outside of the experimental vessel, thus enabling its departure from elastic isotropy to be detected. This work was extended to higher pressures by such workers as **Tait,** but the early work is superseded by that of **Bridgman,** which is now accepted as standard. The liquid was placed in a strong alloy-steel vessel, and the apparent change in volume relative to the vessel measured by the advance of the piston producing the pressure. The pressure itself was measured by the change in resistance of a manganin wire contained in the vessel, and the expansion of the vessel itself could be corrected for by a second set of experiments in which part of the liquid was replaced by a steel plug.

A slightly less direct method of measuring the bulk modulus of a liquid is by measuring the velocity of sound in it. For methods of doing this, the reader is referred to works on sound. Of recent years, the technique of ultrasonics has attained great precision, and reliable measurements can be made in comparatively small samples of liquid. As with a gas, the velocity of sound is proportional to the square root of the *adiabatic* bulk modulus, but the difference between the adiabatic and isothermal bulk moduli can be computed thermodynamically, if the coefficient of expansion is known, and never amounts to more than a few per cent. for a solid or liquid.

We have already referred briefly to measurements of the *tensile strength* of liquids (p. 185).

(*b*) *Viscosity.*—The most direct method of measuring viscosity is by an application of Poiseuille's law, derived on p. 147, and, basically, the experiment is a very simple one. The liquid is driven, by some device for maintaining a constant pressure head, through a capillary, whose internal dimensions are measured and the rate of flow determined by timing and weighing. To obtain really accurate results it is necessary to enclose the apparatus in a thermostat on account of the extreme sensitivity of viscosity to temperature and, in addition, quite a number of corrections are necessary, *e.g.*:—

(i) Hydrodynamic conditions at the entry to and exit from the capillary are not precisely defined, so that there may be some inaccuracy in estimating the true pressure-gradient. This can be overcome, *e.g.* by supplying the capillary with side-tubes, and measuring the pressure-difference between them by a manometer.

(ii) Allowance must be made for the kinetic energy of the emerging liquid.

Commercial "viscometers" are also based on this principle, the procedure being to allow a pipette of known volume to empty itself by flow through a capillary. This enables *relative* values of viscosity to be obtained very rapidly, and, if desired, such an instrument can be calibrated with a liquid of known viscosity and density.

A second method of measuring viscosity is by measuring the damping of the oscillations of a disc suspended horizontally by a torsion wire attached to its centre. This method is due to **Maxwell**, but the theory is very complicated. A method on a similar principle, in which a simple pendulum with a spherical bob was swung in the liquid is due to **Stokes**. Another method is to measure the damping of the torsional oscillations of a hollow sphere suspended on a torsion wire and filled with the liquid. For such simple geometry, solution of the equations of motion is a practical proposition. The reader is referred to Lamb's *Hydrodynamics*.

A third method, very suitable for very viscous liquids such as oil or treacle, is to use an apparatus consisting of two coaxial cylinders, the annular space between which is filled with the liquid [Fig. 75(*a*) and (*b*)], the outer cylinder being rotated at a uniform angular velocity, and the torque on the inner one measured by any of the usual methods, such as a torsion balance, or torsion wire and mirror. Let $\omega(r)$ be the angular velocity of the liquid as a function of the radius r. If the same relation between viscous drag and velocity-gradient held when the streamlines are curved as when they are straight, the force per unit area acting across the cylinder of radius r would be

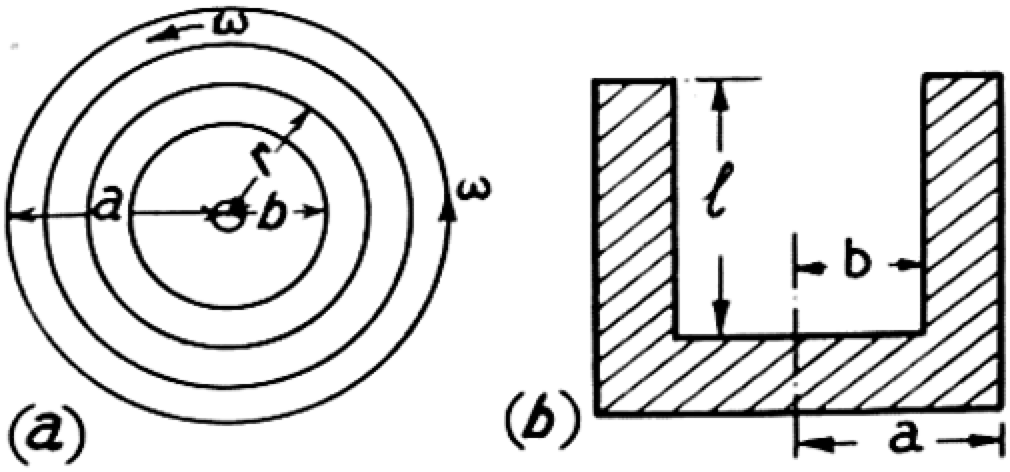


FIG. 75.

$$\eta \frac{\partial v}{\partial r} = \eta \frac{\partial}{\partial r}[r\omega(r)]$$

$$= \eta\left(\omega(r) + r\frac{\partial \omega}{\partial r}\right).$$

However, for curved streamlines the relations between viscous force and velocity distribution are more complicated, and the reader is referred, *e.g.*, to Lamb's *Hydrodynamics* for details. Plainly, the expression just derived cannot be correct, for it would imply that a viscous force was acting even if ω was independent of r, whereas there can be no such drag if the whole liquid is rotating with a uniform angular velocity. It can be shown that the correct expression is obtained from the above by merely omitting the term involving $\omega(r)$. The actual force is thus $\eta r \frac{\partial \omega}{\partial r}$ per unit area, and the total torque acting across a cylindrical surface of radius r and unit length is

$$\eta r \frac{\partial \omega}{\partial r} . 2\pi r . r \quad \dots\dots\dots\dots\dots\dots \quad (3)$$

For steady conditions there can be no torque on the hollow cylinder between r and $r + dr$ because its angular velocity is constant. This means that $r^3 \frac{\partial \omega}{\partial r}$ is constant, that is $\omega = \frac{A}{r^2} + B$, and the constants have to be chosen so that $\omega = \omega_0$ at $r = a$ and vanishes at $r = b$. This gives us

$$\frac{\omega}{\omega_0} = \frac{a^2}{a^2 - b^2}\left(1 - \frac{b^2}{r^2}\right).$$

The torque on either the inner or outer cylinder is then given by

$$G = 2\pi \eta r^3 \frac{\partial \omega}{\partial r} = \frac{4\pi \omega_0 a^2 b^2}{a^2 - b^2} \text{ per unit length} \quad \dots\dots\dots \quad (4)$$

since $r^3 \frac{\partial \omega}{\partial r}$ is constant everywhere between the cylinders.

We have neglected any effect on the bottom of the inner cylinder, but we may expect this contribution to the torque to be almost independent of l, the length of curved surface of the inner cylinder immersed, so that we can eliminate this error by carrying out experiments with two or three different values of l, since equation (4) gives a total torque proportional to l, but the correction is independent of l.

Yet another simple laboratory method of measuring the viscosity of a fairly viscous liquid is to enclose it in a long tube and drop small balls through it and determine their steady velocity of fall when the viscous force balances the effect of gravity. According to Stokes' law, the viscous force on a sphere moving through the liquid with a velocity u is $6\pi\eta a u$, and this has to be equated to the downward force $\frac{4}{3}\pi(\rho_s - \rho_l)a^3 g$, where ρ_s is the density of the solid, and ρ_l that of the liquid. This method is a very rapid one, but has the disadvantage that corrections due to the walls of the vessel are surprisingly large. For a cylindrical vessel, Landenburg arrived at the correcting factor $\left(1 + \frac{2{\cdot}4a}{R}\right)$, where R is the radius of the cylindrical vessel, with the centre of the ball on the axis of the cylinder.

6. Superfluidity

The phenomenon of *superconductivity* in metals has been known for nearly half a century. Its characteristic feature is a *complete* disappearance of the electrical resistance of certain metals at very low temperatures, complete in the sense that a current can be induced in a ring of metal which is accompanied by no measurable e.m.f. and will maintain itself for *several days* at apparently undiminished strength. This phenomenon has necessitated the revision of some of the fundamental ideas of physics, for it seems to indicate that electrons can flow, and form a macroscopic current that is not diminished owing to resistance and other dissipative processes, as happens with any ordinary assembly, such as gas molecules flowing through a tube, in which the energy associated with any *drift* velocity is eventually dissipated as heat.

More recently, evidence for a second type of "superflow" has been obtained. It is found that liquid helium undergoes some kind of transition at about 2·2° K. There is also considerable evidence that below this temperature part at least of the liquid has "superfluid" properties. No one has yet demonstrated a "persistent flow", analogous to the persistent currents found in superconductors, but it is known that the liquid can flow at remarkably small pressure differences, of the order of a fraction of a millimetre head of liquid helium, and also that it can flow through fine channels at speeds far greater than are possible for any other known liquid or gas. Thus, if the very fine particles of jeweller's rouge are compressed to form a pad, ordinary liquids or gases will only penetrate it very slowly under pressures of the order of atmospheres, the dimensions of the channels being of the order of 10^{-6} cm., yet the low temperature modification (usually known as helium II) passes through it very rapidly under pressure heads of a few millimetres of liquid helium.

If the ordinary conception of viscosity were applicable at all, measurements of this kind would point to values of the order of 10^{-10} poise, compared with a value of 10^{-2} poise for water and 3×10^{-5} poise for ordinary liquid helium. On the other hand, if a disc is suspended on a torsion fibre and oscillated in liquid helium II, a definite damping is found, just as it would be in an ordinary liquid. Further experiments on these lines (using a pile of discs instead of a single disc) seemed to indicate that only part of the liquid was being "carried round" by the discs, that is, the effects observed were identical with those to be expected in an ordinary liquid of density smaller than that of the liquid helium. The theory of this experiment is complicated, but the physical reason why this method is capable of measuring effective density as well as viscosity is this: the oscillating disc drags round with it layers of liquid, the velocity of the layers falling off as the distance from the disc increases. The viscous drag *between* these layers results in a *damping* of the oscillation, while the *mass* of these layers results in an effective addition to the *moment of inertia* of the suspended system which can be detected by its effect on the *period* of the oscillation. (This effect is *in addition to* the slight effect on the period produced directly by the damping, as described in the Appendix, p. 289.)

A combination of these two types of experiment leads us to the idea that liquid helium II is capable of flowing in two distinct ways, either as a superfluid which apparently "slips past" solid boundaries and penetrates narrow channels, as would a "perfect fluid" according to classical hydrodynamics, or else as a normal fluid which is capable of interacting with solids and is held back by narrow capillaries. The effective density of the normal fluid can be measured by the oscillating disc experiments, and the remaining mass of the liquid helium must be attributed to the superfluid. One is naturally reluctant to postulate the physical presence of two distinct fluids within something so very simple chemically as liquid helium, but a further set of experiments (on the propagation of temperature differences in liquid helium II) seems to leave no doubt that this "ratio of effective densities" ρ_s/ρ_n is a physically significant quantity, though this crude "two-fluid" conception will, no doubt, be replaced in due course by something more closely related to the atomic structure of the liquid.

The quantity ρ_s/ρ_n varies markedly with temperature and vanishes at the transition temperature. If we have two vessels at different temperatures connected by a fine capillary or slit, the values of ρ_s and ρ_n will be slightly different on the two sides of the slit. Viscosity prevents appreciable flow of the normal fluid but the superfluid flows through the narrow channel until either the temperatures are equalised, or else until a sufficient pressure-difference to stop the flow has been built up. This state of affairs is analogous to what occurs when two solutions of different concentrations are separated by a membrane which can be penetrated only by solvent molecules (see Chapter VIII).

Pressure-differences amounting to several centimetres head of liquid helium can be produced by temperature differences of as little as 0·0001° K. This effect is known as the thermo-mechanical, or "fountain", effect and, within limits, is found to be thermodynamically reversible, that is, a difference in pressure will itself produce a difference in temperature in two vessels connected by a narrow slit. If we make the slit somewhat wider, normal fluid flows through it in the ordinary way in response to the pressure-difference, and it is possible to reach a state of affairs in which normal fluid and superfluid are flowing in opposite directions at equal rates through the slit. This leads to a transport of heat (the heat capacity of the superfluid is believed to be negligible) apparently similar to ordinary conduction of heat, because there is no net flow of liquid between the vessels, but corresponding to a fantastically high apparent thermal conductivity. This was one of the first of the anomalous properties of liquid helium II to be discovered, and was thought to indicate some "super-heat conductivity", but it now seems that it is really a *convection* effect.

This idea is supported by the existence of another phenomenon peculiar to liquid helium II known, rather unhappily, as "second sound". This consists of the propagation, *without appreciable attenuation,* of periodic temperature changes produced, *e.g.*, by applying alternating current to a heating coil. This is something quite distinct from ordinary thermal conduction associated with a periodic variation of temperature for which, however high the numerical value of the thermal conductivity may be, there is always a big attenuation within a distance of the order of the wavelength of the temperature-wave.

On the two-fluid picture, second sound is interpreted as the result of relative motion of the normal fluid and superfluid. Thus ρ_s/ρ_n varies with distance and time, and the value of this quantity at any particular place is a measure of the temperature there. Since the heat content of the superfluid is negligible, second sound may, like the apparent conduction of heat, be regarded as a convection effect due to the motion of the normal fluid. It is possible to set up a "two-fluid hydrodynamics" and from it derive a relation between ρ_s/ρ_n and the observed velocity of propagation of second sound. This is

$$C_2^2 = \frac{\rho_s}{\rho_n} \frac{s^2}{(\partial s/\partial T)_p} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (5)$$

where $T\,(\partial s/\partial T)_p$ is the specific heat at constant pressure, which is a measurable quantity. s, the entropy per unit mass, can be deduced from the specific heat by numerical integration, the Nernst Heat Theorem requiring that $s = 0$ at $T = 0$. Equation (5) then enables us to make a determination of ρ_n/ρ_s, and the values agree closely with those from the oscillating disc experiment, strongly suggesting that this parameter really has a definite physical meaning, in spite of its strangeness.

No satisfactory theory accounting for all these observations yet exists, but it is quite clearly established that there is no hope of an explanation on the basis of classical physics and that quantum mechanics has to be invoked. It has been shown theoretically that certain types of assembly can have quantum states of low energy which are appreciably occupied only below a certain temperature. Equally, if the interactions between the molecules of a liquid are described by means of sound-waves in the manner mentioned above (p. 197) we can calculate the apparent mass "dragged along" by a surface moving tangentially to the liquid, and this mass may, at low temperatures, be less than the actual mass of the liquid. Further, we can account for the fact that liquid helium, unlike all other bodies, does not solidify at any attainable temperature except under pressure. This arises from the fact that, according to quantum mechanics, a given assembly cannot have less than a certain energy (the so-called zero-point energy), even at the absolute zero of temperature. This zero-point energy is greatest for *light* atoms and increases rapidly as the volume of the assembly is diminished, with the net result that a periodic crystal lattice, which is the configuration of lowest energy for all other substances, is the equilibrium one for helium only under a pressure of 25 atmospheres or over. At a pressure of one atmosphere, this effect of the zero-point energy "blows out" the average atomic spacing to a point at which the atomic interactions are insufficiently strong to form a definite crystal lattice, and X-ray studies of both forms of the liquid show a distribution function (p. 182) of practically the same form as that of quite ordinary liquids.

All three of these theoretical results are consequences of quantum, as distinct from classical, mechanics, and it seems certain that they will all have to find a place in any ultimately successful theory.

Yet another circumstance that points clearly to the decisive importance of quantum effects in liquid helium is the fact that superfluid effects are, apparently, absent in the liquid form of the rare helium isotope (of atomic weight 3). Such big differences between the properties of chemically identical atoms are familiar enough in quantum theory, but are almost unintelligible on a "classical" basis.

7. Ultrasonic absorption in liquids

Of recent years great improvements have been made in the technique of measurement of high-frequency sound, and, in particular, it is now possible to make accurate measurements of absorptions and propagation velocities as functions of frequency. The important feature of these measurements is that the measured absorption is often greater, sometimes by a factor as large as 1,000, than the absorption that would be expected on the basis of dissipation of energy by viscosity (the viscosity being measured by such methods as those described above, which we may call "quasi-static" methods, where velocities of flow of the order of a few centimetres per second are involved). However, it is not always easy to be sure that this apparent extra absorption is not instrumental, or due to some cause such as scattering of the sound by solid particles suspended in the liquid, and the early work gave very discordant results.

It can now be stated quite definitely that this discrepancy between "observed" and "calculated" absorptions is a real one, and this result is of great importance, because it shows that the description of a liquid by means of the concepts of compressibility and viscosity is inadequate even on a macroscopic basis. It is *not* possible to argue that the concept of viscosity is breaking down because ultrasonic wavelengths are too short as, even at the highest attainable ultrasonic frequencies ($\sim 10^9$ per second), wavelengths are still very large ($\sim 10^{-4}$ cm.), compared with the distances between molecules. Moreover in certain liquids, such as carbon tetrachloride, the "anomalous" absorption, that cannot be accounted for on the basis of viscosity, increases with rising temperature, whereas the viscosity of all ordinary liquids decreases with rising temperature. Another interesting fact is that the extra absorption is particularly large in liquids such as carbon tetrachloride, carbon disulphide, and benzene, whose valencies are chemically saturated, and which have symmetrical molecules that one might expect to be "well behaved."

We can describe this extra absorption formally in two different ways. The first is by the introduction of the new hydrodynamic concept of "bulk viscosity". It has been known since the time of **Maxwell** and **Stokes** that the conventional equations of hydrodynamics are not the most general possible ones consistent with the assumption: stress is proportional to rate of deformation. It is possible, even in an isotropic liquid, to define a coefficient of viscosity that is related to the bulk modulus in the same way that ordinary shear viscosity is related to the shear modulus of elasticity. For the elastic modulus we have *stress proportional to strain*, for the corresponding viscosity we have *stress proportional to rate of strain*.

Thus, with a finite bulk viscosity, we have a pressure-gradient associated with a uniform expansion or contraction of a liquid. It was shown by **Maxwell** that the kinetic theory of gases (using the ordinary model of attracting rigid spheres) led to a unique relation between shear and bulk viscosities, the bulk viscosity being one-third of the shear viscosity (p. 189), so that a gas *can* in effect be described by a single coefficient, but it has only very recently been realised that this relation need not hold for a liquid, and that, in general, two independent coefficients of viscosity are needed to describe it, just as two *independent* coefficients of elasticity are needed to describe an isotropic elastic substance. (Note that some writers speak of a substance having "zero bulk viscosity", when what they really mean is a substance whose two viscosities obey Maxwell's relation.) Evidently, the excess ultrasonic absorption can always be described *formally* by means of a suitable value of the coefficient of bulk viscosity, but the question of the physical origin of the extra absorption is not solved by doing this. Recently **Eckart** and **Liebermann** have proposed a means, independent of absorption measurements, of measuring this coefficient, and have obtained values for it consistent with the absorption measurements. The reader is referred to the original papers for details of the method and results.

We can divide liquids into four classes, according to the behaviour of the ultrasonic absorption, and this enables us to obtain physical insight into the possible causes of this absorption.

I. Liquids for which absorption is correctly described on the basis of Maxwell's relation. Mercury and liquid argon are in this class. (In mercury there is an extra contribution due to thermal conductivity.)

II. Liquids for which the "excess" absorption varies with temperature in the same way as the viscosity itself. Water and the alcohols seem to belong to this class.

III. Liquids for which the "excess" absorption *increases* with temperature. Benzene, carbon disulphide, and carbon tetrachloride are in this class.

IV. Liquids for which the "excess" absorption could be described only by a viscosity coefficient that depends on the frequency, *e.g.* ethyl acetate and acetic acid.

The conception of a coefficient of second viscosity is of real use only in Class II. Here it seems that the same interaction is responsible both for the ordinary viscosity and for the excess absorption, at least this is a reasonable deduction from the fact that they both vary in the same way with temperature, so that the ratio $\dfrac{\text{observed absorption}}{\text{“classical” absorption}}$ is practically a constant (about 3 for water). It may well be that, in water and the alcohols, the main mechanism responsible for both viscosities is the formation and breaking of "hydrogen bonds"—(a type of loose chemical bond formed, *e.g.*, in an —OH group with an oxygen atom in a neighbouring molecule). An interaction of this type is quite different from those ordinarily assumed in the kinetic theory of gases, one important point of difference is that it is not a central force (p. 8) so that it is not surprising that Maxwell's relation for the two viscosities does not hold here.

In Class III it is clear that, because of the different behaviour as a function of temperature, the mechanism responsible for the "excess" absorption is distinct from that responsible for ordinary viscosity. Attempts have been made to explain this excess absorption on the same lines as the excess absorption of sound in gases such as carbon dioxide. The latter effect has been definitely traced to the inability of the vibrational modes of the molecules themselves to "keep in step" with the ultrasonic oscillations—the result being a dissipation of energy analogous to the losses in a dielectric in an alternating electric field, the dipoles being unable to keep fully in step with the oscillations of the field, see Debye, *Polar Molecules*. The existence of similar effects in liquids can be definitely checked only if ultrasonic frequencies are available comparable with the inverse of the "time of relaxation", *i.e.* the time that would be taken by the degree of freedom concerned to respond to a suddenly applied disturbance. In liquids such times are very short, and it is only in a few cases that an attempt to identify the degree of freedom concerned can be made. If the frequencies available are small compared with the inverse of the time of relaxation, the rate of dissipation of energy is proportional to the square of the frequency, just as it is with ordinary viscosity, and it is not then possible to estimate the time of relaxation. If one has high enough frequencies available, the absorption ceases to be proportional to the square of the frequency, and the shape of the absorption-frequency curve, together with the accompanying dependence of *velocity of propagation* on frequency, then enables the relevant time of relaxation to be estimated—the first step in identifying the mechanism responsible for it. It is probable that such a departure from the frequency squared law of absorption is what is happening in group IV which probably consists of liquids with abnormally long relaxation times.

Although at present at a very tentative stage, it will be seen that ultrasonic investigations promise to give a great deal of information about the interactions between atoms and molecules in the liquid state. For further details, the reader is referred to a review by **Kittel**.

BIBLIOGRAPHY FOR CHAPTER VII

J. Frenkel, *Kinetic Theory of Liquids*. O.U.P. (1946).
C. Kittel, *Rep. Prog. Phys.*, **11**, 205 (1948).
L. N. Liebermann, *Phys. Rev.*, **75**, 1415 (1949).
C. Eckart, *Phys. Rev.*, **73**, 68 (1948).
H. Lamb, *Hydrodynamics*. 6th Edn. C.U.P. (1932).
P. Debye, *Polar Molecules*. Chemical Catalog Co., New York (1929).
R. Fürth, *Proc. Camb. Phil. Soc.* (1941) (a series of papers).
R. H. Fowler and E. A. Guggenheim, *Statistical Thermodynamics*. Revised Edn. C.U.P. (1949).
W. H. Keesom, *Helium*. Elsevier (1941).

R. B. Dingle *et al.*, *Advances in Physics*, Vol. I, No. 2. Taylor and Francis, London (1952).

F. London, *Superfluids*, Vol. II (to be published shortly).

H. N. V. Temperley, *Proc. Phys. Soc. London* (1946-7) (three papers on the behaviour of water under hydrostatic tension).

Examples on Chapter VII

1. Give as many ways as you can by which the "size" of molecules can be estimated experimentally, being careful to point out exactly what property of the molecule is being measured by each type of experiment.

2. Explain what is meant by the distribution function of a liquid. How can information about it be obtained? Can corresponding quantities be defined for a gas or solid?

3. Obtain an expression for the pressure of a van der Waals' gas in descending powers of the volume v.

4. It is desired to design a compact apparatus for routine measurements of viscosity of large numbers of samples of a liquid such as glycerine at room temperature. Design such an apparatus, giving reasons for your choice of dimensions. (Viscosity of glycerine at room temperature 15 C.G.S. units.)

5. Compare and contrast the properties of sound-waves of audible and ultrasonic frequencies in liquids.

6. Write an essay on the theory of liquids.

7. What are the main facts that any ultimately successful theory of liquid helium will have to explain?

CHAPTER VIII

SOLUTIONS AND ELECTROLYTES. OSMOTIC PRESSURE, DIFFUSION, AND RELATED PHENOMENA

1. Introduction

Two of the most striking changes, that may occur when a solid is dissolved in a liquid, are, first that the solution may become a conductor of electricity, even when both the solid and the pure liquid are insulators, secondly, that the solution always tends to dilute itself by attracting more pure solvent to it, if this is possible. The latter effect occurs in various ways, *e.g.* if the solution and pure solvent are placed in contact, diffusion occurs until the concentration is uniform. The same tendency can, however, manifest itself indirectly in many other ways—for example, if a vessel containing solution and a second one containing pure solvent are placed together under a bell-jar, the pure solvent gradually evaporates and the volume of the solution increases by condensation from the vapour. Again, if the solution and solvent are separated from one another by a membrane that will pass solvent but not solute molecules, the solvent molecules try to find their way into the solution, and, if the solution is contained in a closed vessel, a very considerable pressure may be built up in it. This persistant tendency of a solution to dilute itself is known as *osmosis*, and the pressure that has to be applied to the solution in order to stop it taking up more of the pure solvent is called the *osmotic pressure* of the solution.

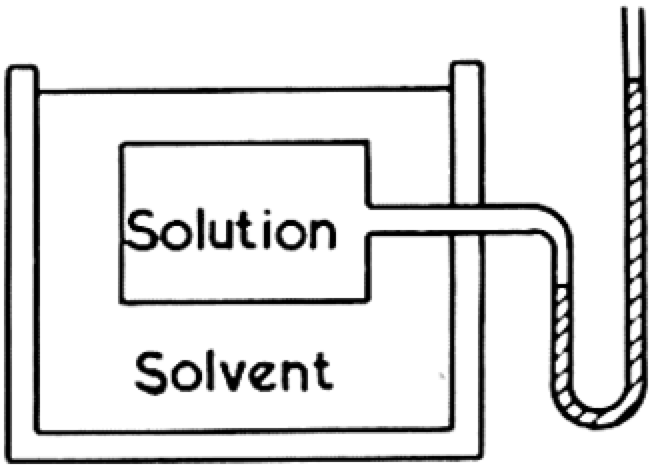


FIG. 76.

2. Measurement of osmotic pressures

The reader is referred to textbooks of physical chemistry for details of the methods of measurement of osmotic pressure. The pressure to be measured may vary from a few millimetres of mercury for a dilute solution or a colloid to many hundreds of atmospheres for a strong solution of a very soluble substance such as sugar. The basic principle is shown in Fig. 76. The solution is contained in a porous pot whose pores are impregnated with some *semi-permeable* material which will pass solvent but not solute molecules. One much used for ordinary solutions is copper ferrocyanide, and the membrane can actually be prepared in the pores of the pot by filling it with a solution of a copper salt and standing the pot in a solution of potassium ferrocyanide. If such a pot, known as a **Pfeffer pot,** is filled with solution and then inserted in pure solvent, the solvent diffuses through the membrane, until the pressure in the pot becomes steady and can be measured by a manometer. For strong solutions there is danger of bursting the pot, and, although special methods have been devised, such as putting the pure solvent *inside* the pot and applying pressure to the solution *outside* until conditions are steady, the measurement of osmotic pressure directly is always difficult. However, there are many possibilities of measuring it indirectly and we shall first derive, by a quite general argument, a relationship between the osmotic pressure of a solution and the vapour pressures of the solution and the pure solvent.

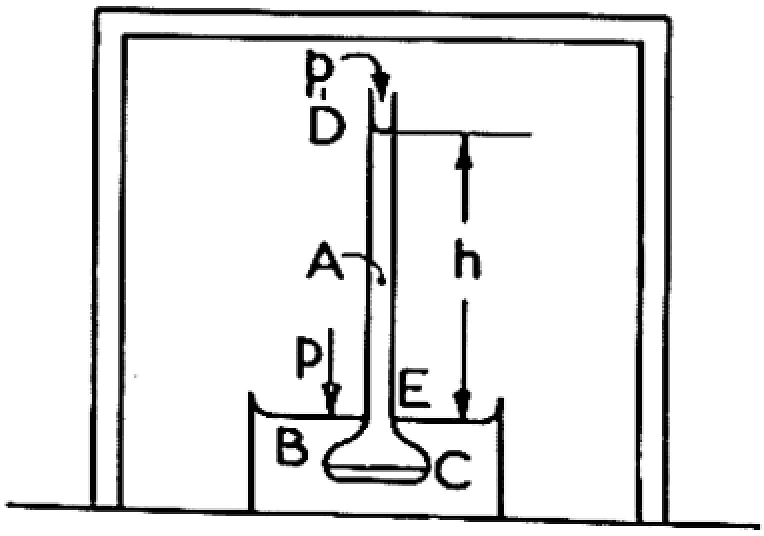


FIG. 77.

Let Fig. 77 represent a state of equilibrium between solution and pure solvent. The solution in vessel A is separated from the pure solvent in vessel B by a semi-permeable membrane C. There will be equilibrium when the pressure corresponding to the length of the column of liquid CD is equal to the osmotic pressure. We may allow the vapour of the solvent access to the top of the column D and the surface of the solvent E without affecting this equilibrium, and we may assume further that there is no other gas present. In such an assembly, no work is done if a molecule of solvent evaporates at D, is condensed in the pure liquid at E, and is finally restored to the solution by passage through the membrane C. In such a cycle, nothing has been changed, and no heat has been gained or lost by the assembly as a whole, so that the work done must be zero. This requires that the pressure p over the surface of the solvent and the pressure p_1 over the surface of the solution, must be related to the osmotic pressure P by the relationship

$$P = p_1 + \rho g h - p \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (1)$$

where ρ is the density of the solution, $p - p_1$ is equal to the difference of vapour pressures over the solvent and solution, and is also equal to the hydrostatic pressure difference associated with a column of vapour of density σ and height h. For a short column of vapour we can write simply $p - p_1 = \sigma g h$, but, more accurately, we ought to take account of the change of density with height (cf. p. 118). We may write

$$dp = -g\sigma dh$$

or, assuming that the vapour obeys Boyle's law, so that $\sigma = \sigma_0 p$, where σ_0 is the density of the vapour at a pressure of one atmosphere,

$$\frac{dp}{p} = -g\sigma_0 dh.$$

Integrating $$\log \frac{p}{p_1} = g\sigma_0 h \quad \ldots\ldots\ldots\ldots (2)$$

Eliminating h between equations (1) and (2) gives us a relation between P and $p - p_1$. If the solutions are dilute enough (2) can be replaced by $p - p_1 = \sigma gh$, so that the relation becomes

$$\frac{p - p_1}{P} = \frac{\sigma}{\rho - \sigma} \quad \ldots\ldots\ldots\ldots (3)$$

which means that, for dilute solutions, the *lowering of vapour pressure is proportional to the osmotic pressure.* For more concentrated solutions we must use equation (2), and, strictly speaking, we ought to correct the theory further to allow for the fact that ρ is not constant but varies with pressure, and also for the fact that vapours do not exactly obey Boyle's law. However, they obey it much more closely than one might expect and these last two refinements of the theory are unnecessary.

Thus the variation of osmotic pressure may be considered known if the lowering of vapour pressure can be measured. At small concentrations they are proportional to one another. It will be noticed that the above argument makes no assumption at all about the mechanism causing osmotic pressure, but merely assumes the *existence* of semipermeable membranes. Thus the argument is a *thermodynamic* one (of the particular kind where no heat exchanges or temperature changes are involved, so that we use only the principle of conservation of energy).

The measurement of changes of vapour pressure can be carried out in various ways described in books on heat. It is possible to do it directly, *e.g.* by introducing a small quantity of the solution into the Torricellian vacuum of a mercury barometer, and comparing the depression of the mercury with that due to the pure liquid, but the most accurate methods are indirect, and depend on measurements of the *elevation of the boiling point* or the *lowering of the freezing point* of the liquid by dissolving the solid in it.

Since a liquid boils when its vapour pressure becomes equal to that of the atmosphere, a lowering of the vapour pressure from p to p_1 must result in a raising of the boiling point by ΔT, where $\Delta T = (p - p_1) . dT/dp$, where dT/dp is the slope of the vapour-pressure curve of the liquid. This is connected with measurable properties of the liquid by the following equation, the celebrated *Clausius-Clapeyron* *equation*, for a derivation of which the reader is referred to textbooks of thermodynamics:—

$$\frac{dp}{dT} = \frac{l}{T\,(v_{\text{vap}} - v_{\text{liq}})} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (4)$$

where l is the latent heat of evaporation per unit mass, v_{vap} and v_{liq} are the volumes of unit mass of the liquid and vapour respectively.

The connection between the lowering of the vapour pressure and the depression of the freezing point is slightly less direct, but may be obtained in the following way. Nearly all *dilute* solutions when frozen deposit *pure* solid solvent at first. Freezing thus occurs when the vapour pressure of the solid and that of the solution are the same, so that both the solid and the solution are in equilibrium with the same vapour, and therefore with one another. The vapour pressure of the solution is less than that of the pure solvent, consequently the solid is in equilibrium with the solution at a lower temperature than it would be with the pure solvent. The change in the vapour pressure of a solid with temperature is also governed by equation (4), where l is now the *heat of sublimation*, so that again we can compute, in terms of physically measurable quantities, the relation between lowering of vapour pressure and depression of freezing point, that is the temperature difference corresponding, on the vapour-pressure curve of the solid, to a pressure-difference of $p - p_1$. The *changes* in boiling point and freezing point can be measured with satisfactory accuracy by a thermometer such as Beckmann's, the chief difficulties in temperature measurements arising when *absolute* measurements of temperature are required.

3. Theory of osmotic pressure

The experimental facts and theoretical reasoning that we have outlined in the last article tend to give the impression that osmotic phenomena are really the manifestation of attractive forces between the solute and solvent molecules. Indeed, we know for certain that many substances combine chemically with water when dissolved, and the existence of strong attractive forces between solvent and solute is probably very common. Nevertheless, the view that we have just outlined is false. Osmotic pressure would exist even in a so-called "perfect solution", that is to say a solution in which the interactions do not distinguish between solvent and solute molecules, so that there is no heat of solution. We shall first develop the laws of perfect solutions, and shall then point out in what ways the existence of a heat of solution modifies these laws.

The true explanation of osmotic pressure is that it is due to the *thermal agitation of the solute molecules*. The solvent molecules can pass through the semi-permeable membrane, from which it follows that, in equilibrium, the mean numbers striking the membrane on each side must be the same, this argument being familiar in the kinetic theory of gases. However, the solute molecules also strike the membrane but are turned back by it, and the resulting destruction of momentum results in an excess pressure on the solution side of the membrane which has no counterpart on the other side. Thus it is only possible to have equilibrium if the *total* pressure in the solution is greater than that in the solvent. One can put the argument in another way, which may help to show more clearly why solvent tries to flow into the solution. Let us suppose that we start an experiment with the total pressures equal on both sides of a semi-permeable membrane. On the solution side of the membrane this pressure is due partly to the solute molecules striking the membrane, and the remainder to the solvent molecules. On the pure liquid side we have only the effect of the solvent molecules. If the *total* pressures on both sides are known to be equal we must conclude that more solvent molecules are striking the membrane from the pure liquid side than from the solution side, and this must result in a net transfer of liquid through the membrane.

If the solution is *dilute*, the solute molecules will be surrounded by solvent molecules, and will encounter one another comparatively rarely, so that their interaction with one another is slight, and they may thus be thought of as moving about independently. In these circumstances we can use Boltzmann's principle to show that the mean kinetic energy of the solute molecules is $\frac{3}{2}kT$, even in the presence of an external field of force (p. 118). If their interactions with one another are negligible, we can take a second step and infer the same connection between pressure and mean square velocity that we deduced for a perfect gas by considering the rate of destruction of momentum at the boundary, that is, we may conclude that

$$P = \tfrac{1}{3}Nm\,\bar{c}^2 = NkT \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (5)$$

where P is the pressure due to the solute molecules alone, m is the mass of each, and N is the number per unit volume. Thus the osmotic pressure should, for dilute solutions, obey the same law of variation with temperature as the pressure of a perfect gas. This result was predicted by **van't Hoff** by a somewhat different argument and is experimentally confirmed for dilute solutions.

An equation such as (5) cannot be expected to describe the osmotic pressure of a *concentrated* solution, where interactions between solute molecules become important. The pressure in a pure liquid is very much less than that suggested by an equation such as (5), because the term NkT, representing the effect of thermal agitation, is very nearly balanced by the intermolecular attraction (described roughly by the van der Waals' term a/v^2). For this reason, the *change* in the pressure resulting from the presence of solute molecules, and a consequent increase in the term NkT, may be comparable with the pressure in the pure liquid even though the number of added molecules is only a small fraction of N. If the solute molecules interact among themselves, relation (5) ceases to be true, just as the analogous equation for a gas had to be modified to allow for molecular interactions. It is very important to realise that it is the second step in the argument given in the last paragraph that breaks down, not the first. The result that the mean *kinetic energy* of the solute molecules is $\frac{3}{2}kT$ still holds, even when strong interactions between the molecules exist (provided that these interactions depend only on the relative *positions* of the molecules rather than on their velocities. This is probably nearly true, except for very dense assemblies). It is the second step of the argument that breaks down, because, when we consider the destruction of momentum at the membrane, this is partly effected by the forces exerted by the membrane on the molecules, the "pressure", and partly by the interaction forces between the solute molecules themselves. On p. 190 we saw how the latter effect could, for the gas, be allowed for to a first approximation, leading to an equation of state of van der Waals' or similar type. As we have seen (p. 195), the order of magnitude of the van der Waals' constants in a gas can be calculated by quite a number of independent methods. A number of similar equations have, in fact, been proposed for the osmotic pressure of strong solutions, but, as we do not usually have any independent knowledge of the forces between the solute molecules, it is hardly possible to correlate the "osmotic equation of state" with other physical properties. One important exception to this rather sweeping statement must be made—the Debye-Hückel theory of "strong" electrolytes. Here the solute molecules break up into charged ions, whose law of interaction is precisely known. The effect of this interaction on the osmotic pressure can be allowed for, and, at all events for dilute solutions, it seems to be more important than the effect of the short-range forces of interaction between ordinary uncharged molecules. We shall return to this point later (p. 223), but we now consider the physical implications of equation (5).

The great physical and chemical importance of equation (5) is that, if we can measure the osmotic pressure P, or infer it by one of the indirect methods described above, we can at once determine N, the *number* of molecules of solute per unit volume, and this is irrespective of whether they move about independently, or carry solvent molecules with them, that is to say that, since we know the total weight of the solute, *we can determine its effective molecular weight in solution.* Speaking more precisely, we can determine whether, when we dissolve a solid in a liquid, it really does break up into molecules, and, if so, their size. Thus, if we dissolve a pure element in a liquid, we have no means of knowing *a priori* the number of atoms that will be bonded together to form a molecule, but this method will easily tell us. Again, if we have a preparation of a protein, or similar very large molecule, whose molecular weight is unknown, knowledge of the osmotic pressure would enable us to determine its molecular weight. Again, in colloids, see Art. 4, the typical particle contains a great many molecules, and the osmotic pressure is consequently very small, but this is a method by which their size can be estimated just when other methods, such as microscopic observation, fail because the particles are too small.

The "effective molecular weight", as inferred from osmotic pressure, or one of the quantities that depend on it, enables us to determine the size of the "unit" that makes up a solution, just as the determination of vapour density enables us to determine the effective molecular weight. It sometimes happens that a substance when vaporised has a smaller molecular weight than when dissolved in a liquid; this may mean that, in the vapour, the molecules act as distinct units, whereas, in solution, residual chemical forces result in two or more molecules joining together, processes sometimes known as *association* or *polymerisation.*

Another possibility is that the molecules of solute may interact chemically with molecules of solvent, which become tightly bound to the molecules of solute. This increases the effective mass of the molecules of solute, but equation (5) shows that osmotic pressure is independent of the *mass* of a molecule of solute, and depends only on N/V, the number of *discrete units* per unit volume.

4. The theory of solutions

We can divide solutions into three classes, according to the behaviour of the osmotic pressure, and the corresponding apparent value of N/V.

(*a*) N/V is *less* than the value expected from the molecular formula determined by other means. This may mean that either:—

(i) Association or polymerisation has occurred.

(ii) The particles present are very much larger than true molecules so that we really have a *suspension.* Such suspensions, often known as *colloids*, because *gums* are typical examples, are often quite stable. The suspensions are prevented from settling by a variety of causes, the chief of which are the existence of *electric* *charges* on the particles and the *Brownian movement* (p. 120) which is the result of thermal agitation.

(*b*) N/V is *equal* to the value expected from the molecular formula. Such solutions are usually called *undissociated* solutions. A typical example is sugar dissolved in water.

(*c*) N/V is *greater* than the value expected from the molecular formula. This indicates that the molecule has *broken up* in solution. This usually occurs when acids, bases, or salts are dissolved in water, in which case the solution becomes capable of conducting electricity. This is caused by the breaking-up of the molecules into charged particles or ions (which recombine if the solvent is evaporated away). This process is called *electrolytic dissociation*. Water is not the *only* solvent in which this phenomenon occurs, but it is far less common in other solvents. One reason for this is certainly the very high *dielectric constant* of water, nearly 80 as compared with about 2 or 3 for ordinary liquids. We shall discuss electrolytes further below (p. 221).

We consider now some effects of interactions between the molecules. For dilute solutions we have seen that the relative lowering of vapour pressure is proportional to the osmotic pressure, which is, in turn, proportional to the concentration. This relationship between lowering of vapour pressure and concentration is known as *Raoult's Law*. If such a law is to hold over the whole range of concentrations, it must take the form

$$p_A = p_A{}^0 \frac{N_A}{N_A + N_B} \quad \cdots\cdots\cdots\cdots\cdots\cdots \quad (6)$$

where p_A is the partial pressure associated with the component A, $p_A{}^0$ is the vapour pressure of the pure liquid A and N_A, N_B are the numbers of molecules of A and B present per unit volume. It can be shown, by statistical mechanical methods, that a law of this kind should hold if the molecules of A and B are so alike that the *interaction energies between like and unlike molecules are the same*. A few pairs of chemically similar substances, *e.g.* benzene and cyclohexane, are known for which Raoult's law does hold accurately for all concentrations, but it is usually found to hold only for dilute solutions. (Our indirect discussion, via the osmotic pressure, also indicated that, in general, Raoult's law is only likely to hold in the limit of dilute solutions for which osmotic pressure is proportional to concentration.)

Solutions for which the interaction energies between components give no preference between like and unlike pairs of molecules are known as *perfect solutions* and should obey Raoult's law. Departures from Raoult's law can be caused by interaction energies tending to bring either like or unlike molecules together. The *sign* of the deviation from Raoult's law can be easily determined by a simple physical argument, but only rough quantitative treatments of solutions with a finite "energy of mixing" are available at present. (The reader is referred to treatises on statistical mechanics for details.) Suppose first that the interaction energies are such as to favour like molecules being close together. We saw above that the osmotic pressure of a perfect solution was a measure of the extra pressure associated with the thermal agitation of the solute molecules. If now these are, on the whole, *attracted* by one another, their tendency to escape from the solution into a region occupied by pure solvent is lessened and therefore the osmotic pressure will be *less* than if the solution were perfect, a result that may be expected for all concentrations, though the effect will be small for dilute solutions. The osmotic pressure is proportional, not directly to the vapour pressure of either component, but to the *relative lowering* of the vapour pressure of the solvent due to the presence of the solute. Thus, for a perfect solution we should have from equation (6), if we reckon component B as the solute:

$$\begin{array}{l}\text{Relative lowering of}\\ \text{vapour pressure of } A\end{array} = \frac{p_A{}^\circ - p_A}{p_A{}^\circ} = \frac{N_B}{N_A + N_B} \quad \ldots\ldots\ (7)$$

For the actual solution we have the osmotic pressure, and therefore the relative *lowering* of vapour pressure, both *less* than those for perfect solutions, which means that

$$p_A{}^\circ - p_A < \frac{p_A{}^\circ N_B}{N_A + N_B},$$

i.e.

$$p_A > \frac{p_A{}^\circ N_A}{N_A + N_B},$$

that is to say, that the actual vapour pressure is *greater* than that given by Raoult's law, but the osmotic pressure is *less* than the perfect solution value. This state of affairs is called a positive departure from Raoult's law. An exactly similar argument can be used for the case when the intermolecular forces are largest between *unlike* molecules. We shall now get a greater osmotic pressure, but a smaller vapour pressure of the solvent, than we should get for a perfect solution. This is called a *negative* departure from Raoult's law. These results are in the directions we should naturally expect. A strong attraction between solvent and solute molecules must, as we have just seen, *increase* the osmotic pressure, but osmotic pressure is present even in a perfect solution. Equally obviously, such an attraction opposes the tendency of solvent molecules to escape from the solution.

5. Solubilities

Since many of the theoretical treatments of solutions break down when the solutions become too concentrated, it is to be expected that only semi-quantitative theories of solubility curves are possible, but it seems worth while to summarise a few of the ideas that have been put forward. The reader is referred to treatises on chemical thermodynamics and statistical mechanics for details.

(*a*) *Gases in liquids.*—We may regard such a solution as a limiting case of a liquid mixture. A liquid of low boiling point would have a very large vapour pressure at ordinary temperatures. Equation (6) shows that, if the solution is perfect, the concentration should be proportional to the partial pressure of the gas (Henry's law) and also that, at a given pressure, a gas with a low boiling point should be less soluble than one with a high one because of the larger vapour pressure that would be associated with the more "permanent" gas. This tentative conclusion is borne out in practice, though the idea of "vapour pressure" becomes meaningless above the critical temperature of the gas.

(*b*) *Liquids in liquids.*—A very simple argument shows that two liquids that form a "perfect solution" should mix in all proportions. To see this, consider grains of sand of two different colours shaken together. Simple considerations of probability show that the number of possible configurations corresponding to complete mixing is greater by far than the number of possible configurations where there is known to be an appreciable separation of the colours. Thus, a mixed state is overwhelmingly probable. The same is true of two liquids forming a perfect solution, and is true *a fortiori* if the interaction tends to bring unlike molecules together. It is only when *like* molecules *attract* one another that any question of a limited solubility arises. For a sufficiently strong interaction energy, the tendency of the two types of molecule to separate can overcome the tendency of thermal agitation to keep them mixed, and we get a separation into two phases, in the manner shown in Fig. 72. A mixing curve of this general type has been obtained by various crude theoretical treatments, but partially miscible liquids show a great variety of behaviour, for example, some pairs of liquids are known for which the critical mixing curve resembles that of Fig. 72 *turned upside down*, the two-phase region being above it. Such a pair is triethylamine and water. Such mixing curves are not yet understood.

(*c*) *Solids in liquids.*—The most common type of solubility curve of a solid is one that is convex to the temperature axis, so that it is of the same general shape as the left-hand half of the curve shown in Fig. 72, but here again the theory is at a very tentative stage, and we do not really know how to take account of such factors as differing sizes of solvent and solute molecules. Some substances have an almost straight solubility curve, others, such as anhydrous sodium sulphate, show a decrease in solubility as the temperature rises.

(*d*) *Solids in solids.*—All three possible types of behaviour already mentioned are represented here. Corresponding to *perfect solutions* we have pairs of metals such as silver and gold which alloy in all proportions, and pairs of salts such as ferrous and cupric sulphates, which crystallise in the same form with practically identical lattices and are capable of forming "mixed crystals". Corresponding to interactions that favour *like* neighbours we have pairs of metals which will only alloy with one another in certain proportions, or not at all, while, corresponding to interactions that favour *unlike* neighbours, we can have a phenomenon that has no known analogue in liquid solutions, namely the formation of a "superlattice". A typical example is the alloying of copper and zinc to form β-brass. This is a body-centred cubic lattice, and, above a certain temperature, the copper and zinc atoms are arranged on the lattice at random, while below this temperature there is a preponderance of copper atoms on the cube corners and of zinc atoms on the cube centres, this change in structure being detectable by X-ray analysis and by the effect on electrical resistance.

(*e*) *Gases in solids.* The best-known example is hydrogen in palladium, but many examples of gases penetrating into the crystal lattices of solids are known. This is quite distinct from *adsorption* (p. 267), which occurs only on the surface of the solid.

6. Electrolytes

The correct interpretation of the flow of current through an electrolyte seems to have been due to **Faraday,** who realised that the current was carried by charged ions, which were discharged on reaching the electrodes. However, he never seems to have stated definitely that the ions were present in the solution even when no current was flowing. The evidence for such a conclusion is two-fold. In the first place, if we eliminate the effect of secondary phenomena at the electrodes themselves, the current is always strictly proportional to the potential gradient, even when the latter is very small, so no work seems to be required to *create* the ions, secondly, we have the evidence from osmotic and related phenomena, briefly referred to above (p. 215). The depression of the freezing point and the raising of the boiling point of water, when *e.g.* common salt is dissolved in it, are both nearly twice as great as one would expect on the basis that the molecules of salt were undissociated. If, however, it is supposed that the salt is very largely dissociated into ions, each of which is behaving as an independent particle, the enhancement of osmotic pressure can be understood, as can the increased lowering of vapour pressure, from which follow the enhanced effects on boiling and freezing points.

The fact that dissociation occurs in water but not usually in other solvents can be traced to the high dielectric constant of water, which lessens the electrostatic energy required to separate the ions. It is now natural to ask whether dissociation into ions is partial or virtually complete, and it is found that dissociating solutions can be roughly divided into "weak" and "strong" electrolytes, strong electrolytes being those for which the dissociation is complete while "weak" electrolytes are those for which the degree of dissociation depends on the dilution and only approaches completeness for very great dilutions. This behaviour of "weak" electrolytes can be fairly well described by *Ostwald's dilution law*, which treats the process of dissociation as if it were governed by chemical equilibrium between the ions and undissociated molecules. If a solution is made up of C molecules per unit volume and a fraction α of them are dissociated, we can determine the equilibrium condition from the following considerations. The number of molecules that dissociate in unit time is approximately proportional to the number present, that is to say to $C(1-\alpha)$, whereas the number of ions that join up to form a molecule is proportional to the number of collisions between positive and negative ions, that is to the product of the numbers of ions of each kind present, or to $C\alpha . C\alpha$. In equilibrium the numbers of molecules that dissociate and re-form must be equal, so that we have

$$C^2\alpha^2 = KC(1-\alpha) \quad \dots\dots\dots\dots\dots\dots \quad (8)$$

where K is a constant. The student of physical chemistry will recognise equation (8) as a particular case of the *law of mass action* governing chemical equilibria. If, as is usually the case for weak electrolytes, α is small compared with unity, equation (8) tells us that the *degree of dissociation should vary inversely as the square root of the concentration*, which is Ostwald's dilution law. If α is small, it cannot be inferred accurately from osmotic pressures or allied effects, but measurements of the electrolytic conductivity per mole of solute at a number of different dilutions enable us to determine α. At very great dilution the conductivity per mole approaches a limiting value, characteristic of complete dissociation, so that the ratio λ/λ_∞ of actual *conductivity per mole of solute* to the conductivity per mole which we should get at very high dilutions, should be a direct measure of α. (This involves the concealed assumption that the *force resisting the motion* of an ion through the liquid is the same at all sufficiently high dilutions, but this assumption is a very reasonable one.)

As a by-product, this method gives us information on the drift velocity of the ions under the influence of a field. It by no means follows that both positive and negative ions are moving at the same speed. If they are not, the result is not any actual separation of the ions (which would at once bring large electrostatic fields into being) but the appearance of a *gradient of concentration* in the electrolyte—the concentration increasing in the neighbourhood of the electrode towards which the faster-moving ion is travelling, and decreasing near the other electrode. **Hittorf** showed that measurement of this effect (by volumetric analysis of various regions of the electrolyte) enables one to infer the *relative* speeds of the two ions, while conductivity measurements give the *sum* of the speeds of the two ions if they have equal valencies (or an appropriate average of the two speeds if the valencies are different). In this way we can arrive at the absolute speeds of various types of ion as a function of the field. Like the drift velocities of electrons in metals, ionic drift velocities are surprisingly small, of the order of a small fraction of a centimetre per second for ordinary voltages. The rate of drift can also be inferred from measurements of diffusion coefficients for electrolytes, and there is satisfactory numerical agreement. The steady rate of drift in a known applied field enables us to compute also the *force* resisting the motion of an ion through the liquid, and the application of Stokes' law (p. 154) suggests that the ion carries a large number of water molecules with it. For details of these measurements the reader is referred to works on physical chemistry.

Equation (8) enables us to make a definite prediction about the behaviour of the osmotic pressure of an electrolyte as the solution is made very dilute. We confine ourselves for simplicity to solutes that form only two ions per molecule, the work being readily generalised to more complicated molecules. We define the *osmotic coefficient* β of the solution as the ratio of the actual osmotic pressure to the value that it would have if the solute were completely dissociated. This is connected with a, the *degree of dissociation* as follows. We have in unit volume of solution of concentration $C(1-a)$ undissociated molecules, Ca positive ions, and Ca negative ions. Each of these contributes to the osmotic pressure equally, so, by equation (5), we should have

$$P = [C(1-a) + 2Ca]\,kT = C(1+a)\,kT$$

whereas, if dissociation were complete, we should have $P = 2CkT$, so that β should be equal to $(1+a)/2$ in dilute solutions. By equation (8) a is nearly equal to 1 at high dilutions, and, in this region, equation (8) shows us that $1-a$ should be proportional to C. Since $\beta = (1+a)/2$, this means that $1 - \beta$ *should be proportional to the concentration at high dilutions.* For molecules forming more than two ions, the appropriate generalisation of the argument shows that $1 - \beta$ should, for C small, be proportional to $C^{\nu-1}$, where ν is the number of ions per molecule.

This prediction can be checked by measurements of osmotic pressure or related quantities, and it soon appears that, while there are some electrolytes that obey these predictions, there are also many which do not. We now believe that electrolytes can be divided into "weak" electrolytes, for which the above prediction of the variation of $1 - \beta$ with $C^{\nu-1}$ is verified, and "strong" electrolytes, for which $1 - \beta$ is, in the limit, proportional to $C^{\frac{1}{2}}$, which cannot be accounted for on the above basis. Generally speaking "weak" electrolytes consist of organic acids and bases and their salts, also sulphides and carbonates and some ammonium salts. "Strong" electrolytes comprise the inorganic acids, caustic soda and potash and their salts. If a "strong" acid forms a salt with a "weak" base or vice versa the result is usually a "strong" electrolyte.

Modern theory holds that the division of electrolytes into "weak" and "strong" stems from the fact that strong electrolytes are virtually completely dissociated into ions even in concentrated solution. The law of mass action is now irrelevant, and some other explanation must be sought for the departure of the osmotic pressure from the "ideal" value corresponding to complete dissociation. This explanation **Debye** and **Hückel** found by considering the electrostatic interaction between the ions. We now give an outline of their theory. We may calculate the expected osmotic pressure of a strong electrolyte by consideration of the following set of processes. We suppose our solution enclosed in a cylinder, by a piston fitted with a semi-permeable membrane, so that, if we have pure solvent at the far side of the piston we can change the dilution by varying the volume of the solution.

(i) We pull out the piston so as to make the solution very dilute. The work done by the osmotic pressure is w_1, and we may write

$$w_1 = \int_{v_1}^{v_2} P\,dV,$$

where P is the osmotic pressure.

(ii) While the solution is dilute, we remove the charges from the ions. This is an ideal process of a kind often used in thermodynamic arguments. There is, of course, no actual means by which it can be done. If we suppose that there are N ions of each sign present altogether, and that each one is electrically equivalent to a sphere of radius b with a charge ϵ in a medium of dielectric constant D, then the total potential energy of the charges, which is equal to the work done in removing or neutralising them is $N\epsilon^2/bD$, since we may think of each ion as a charged sphere of capacity bD. (We can neglect other types of interactions between the solute molecules if the dilution is great enough).

(iii) We compress the solution of discharged ions back to their original volume. Since we have removed the charges, and we are agreeing to neglect other interactions, the osmotic pressure takes its "ideal" value $2NkT/V$ and the work done is $w_2 = \int_{v_1}^{v_2} \frac{2NkT}{V} dV$.

(iv) We now gradually restore the charges in such a way that, at a given instant, each ion has the same fraction of its final charge, the ions being supposed to be in the volume v_1. The work necessary to recharge an ion may be divided into two parts, the work done in progressively placing the charge ϵ on the ion against the field of the charge already there, and the work done in bringing up the charge ϵ against the field of the other ions. The first part just cancels the term $N\epsilon^2/bD$ that we have already calculated for the dilute solution (we can probably neglect any variation of D with concentration), the second part is simply $\int_0^{\epsilon} \phi(\eta)\, d\eta$, where $\phi(\eta)$ is the electrostatic potential at the position of any ion due to the presence of the others when the hypothetical charging process has reached the stage at which each ion has the charge η.

The solution has been restored to its original state by the hypothetical cycle of operations (i)-(iv) and the total work done must be zero. The two terms $N\epsilon^2/bD$ cancel, and we have

$$\int_{v_1}^{v_2} P\,dV = \int_{v_1}^{v_2} \frac{2NkT}{V}\,dV - 2N\int_0^{\epsilon} \phi(\eta)\,d\eta.$$

If we suppose ϕ known as a function of volume, this equation is equivalent to

$$P_{(V=v_1)} = \frac{2NkT}{v_1} - 2N\frac{\partial}{\partial V}\int_0^{\epsilon} \phi(\eta)\,d\eta \quad \dots\dots\dots\dots \quad (9)$$

Our calculation thus reduces to the problem of determining ϕ when the charge on each ion is $\pm\eta$. Plainly, the average of ϕ over the whole volume is zero, but there will be a preponderance of positive ions near a negative ion and a preponderance of negative ions near a positive one, so that ϕ and η will change sign together. In what follows we drop the suffix ($_1$) from v_1. We shall determine ϕ in the neighbourhood of a positive ion, and the potential in the neighbourhood of a negative ion is $-\phi$. Consider a small volume dV in the neighbourhood of a positive ion where the potential is ϕ. The number of positive ions within this volume, when the charging of each has reached the stage η, is, by Boltzmann's expression, $Ae^{-\frac{\eta\phi}{kT}}\,dV$, while the number of negative ions is $Ae^{+\frac{\eta\phi}{kT}}\,dV$, A being a constant which we must take equal to N/V, since, at high temperatures, the number of positive or negative ions would be N/V per unit volume. Thus, the mean charge density ρ in the neighbourhood of a positive ion is

$$\rho = -\frac{N\eta}{V}\left(e^{+\frac{\eta\phi}{kT}} - e^{-\frac{\eta\phi}{kT}}\right) = -\frac{2N\eta}{V}\sinh\left(\frac{\eta\phi}{kT}\right).$$

Now, by Poisson's equation (see Jeans, *Electricity and Magnetism*) the electrostatic potential and charge density are related by the equation

$$\frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} + \frac{\partial^2\phi}{\partial z^2} = -\frac{4\pi\rho}{D}.$$

Substituting for ρ we get the following equation for ϕ:—

$$\frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} + \frac{\partial^2\phi}{\partial z^2} = +\frac{8\pi N\eta^2\phi}{DkTV} \quad \ldots\ldots\ldots\ldots (10)$$

since we can replace $\sinh(\eta\phi/kT)$ by $\eta\phi/kT$ when ϕ is not too large, that is, the solution not too concentrated.

We are interested in a solution of equation (10) that is spherically symmetrical about each ion, *i.e.* is a function of r only. If we express Poisson's equation in spherical polar coordinates and drop the terms depending on the angular coordinates we get

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d\phi}{dr}\right) = \frac{\phi}{l^2} \quad \ldots\ldots\ldots\ldots\ldots\ldots (11)$$

where l is a characteristic length given by $l = (DkTV/8\pi N\eta^2)^{\frac{1}{2}}$. The general solution of equation (11) is

$$\phi = \frac{Ae^{-r/l}}{r} + \frac{Be^{+r/l}}{r} \quad \ldots\ldots\ldots\ldots\ldots\ldots (12)$$

Since the potential must be finite for r large, we must have $B = 0$.

Expression (12) is the potential at large distances from the selected ion, and will hold at distances from it large enough for it to be possible to treat the remaining ions as a continuous distribution of charge. We may choose a sphere of radius a, the "effective radius" of the ion, within which another ion is unlikely to be found. This distance is not less than the atomic radius, but may well be considerably greater. Inside $r = a$ the potential at each stage of the charging process is mainly that due to the single charge η in a medium of dielectric constant D.

This potential must satisfy Laplace's equation, since other charges are unlikely to be within this sphere. The most general potential possible is $\phi_1 = \eta/Dr + C$, where C is arbitrary. We can now choose A and C so that $\phi = \phi_1$ and $\partial\phi/\partial r = \partial\phi_1/\partial r$ at $r = a$, that is, the potential and the field are both continuous. These conditions give

$$A = \frac{\eta e^{a/l}}{D\,(1 + a/l)}, \qquad C = -\frac{\eta/l}{D\,(1 + a/l)} \quad \dots\dots\dots \quad (13)$$

C represents the potential at the position of the ion considered *due to the charges of the other ions*, and is thus the $\phi\,(\eta)$ to be inserted in equation (9). (The term η/Dr is the potential of the charge on the ion itself, whose effect has already been allowed for in one of the mathematically cancelling terms $N\epsilon^2/bD$.) For small concentrations, we shall always have a/l small and so we may write

$$\phi\,(\eta) = C = -\frac{\eta}{lD} = -\eta^2\left(\frac{8\pi N}{VkTD^3}\right)^{\frac{1}{2}}$$

Substituting in equation (9) we get

$$P = \frac{2NkT}{V} - 2N\frac{\partial}{\partial V}\int_0^{\epsilon}\phi\,(\eta)\,d\eta$$

$$= \frac{2NkT}{V} + 2N\frac{\partial}{\partial V}\frac{\epsilon^3}{3}\left(\frac{8\pi N}{kTVD^3}\right)^{\frac{1}{2}}$$

$$= \frac{2NkT}{V}\left(1 - \frac{\epsilon^3}{6kT}\left(\frac{8\pi N}{kTD^3V}\right)^{\frac{1}{2}}\right) \quad \dots\dots\dots\dots \quad (14)$$

Thus, this relation shows that $1 - \beta$ should be proportional to $(N/V)^{\frac{1}{2}}$ as is in fact observed at low concentrations, the numerical agreement also being satisfactory. The physical interpretation of the characteristic length l is the radius of the "atmosphere" within which the effect of an ion makes itself felt. Outside this its effect is "screened off" by a local preponderance of ions of the opposite sign, this effect being described by the factor $e^{-r/l}$.

Equation (14) is the main result of the Debye-Hückel theory. Some attempts to improve some of the approximations have only served to raise fresh difficulties. The reader is referred to Fowler and Guggenheim, *Statistical Thermodynamics*, for a very full discussion of this very interesting field. This notion of an ionic atmosphere also disposes of another difficulty in the theory of strong electrolytes, namely why their conductivities do not quite correspond with those expected on the basis of complete dissociation, but are always somewhat smaller. It will be seen that each ion "carries with it" an "atmosphere" of ions preponderantly of the opposite sign, and this causes the measured conductivity to fall below the "ideal" value. At very high applied fields the effect of the "atmospheres" is overcome, and the conductivity approaches the "ideal" value (*Wien effect*).

The anomalies in "strong" electrolytes can therefore be accounted for on the basis of the electrostatic interaction between the ions and there is, therefore, no reason to doubt the prevailing view that such molecules are completely dissociated in solution. The application of the law of mass action to *weak* electrolytes also seems entirely justified.

We can see the great difference between the increases in conductivity per mole as "strong" and "weak" electrolytes are diluted. For "weak" electrolytes more and more ions become available but their interactions probably do not greatly matter, while for "strong" electrolytes the number of ions is unchanged, but their motion in the electric field is less hindered by the effect of the "atmospheres".

7. Diffusion

Since osmotic effects are associated with surprisingly high pressure differences, it is natural to argue that large forces may be associated with any non-uniformity of concentration of a solution. The process of diffusion may be thought of as the motion of the solute molecules in response to the differences of osmotic pressure. Consider two neighbouring volumes, each of volume V, of solution of a non-electrolyte of concentrations n_1 and n_2 molecules per unit volume. We want the work done per molecule of solute transferred from the higher to the lower concentration. We imagine first a volume dV of pure *solvent* transferred across the boundary. The work done in removing it from the solution of concentration n_1 is P_1dV, and the work gained in adding it to the solution of concentration n_2 is P_2dV, where P_1 and P_2 are the osmotic pressures associated with the concentrations n_1, n_2 respectively. The net work done is thus $(P_1 - P_2)\,dV$. This follows from the definition of osmotic pressure. We may, if we wish, think of the two volumes of liquid as separated by a semi-permeable membrane. The net effect of the transfer is to change the concentrations in the two volumes to $n_1V/(V - dV)$ and $n_2V/(V + dV)$. To the first order of small quantities these are equal to $n_1(1 + dV/V)$ and $n_2(1 - dV/V)$ and, if n_1 and n_2 are both nearly equal to n, the same changes of concentration could have been produced by transferring ndV molecules of *solute* from the solution of concentration n_2 to that of concentration n_1. Since the osmotic pressure is proportional to n, for dilute solutions, we may conclude that the work done in transferring ndV molecules of solute from a concentration n_2 to a concentration n_1 is

$$(P_1 - P_2)\,dV \simeq (n_1 - n_2)\,kT\,dV$$

by van't Hoff's equation (5). If now we have a solution with a gradient of concentration $\partial n/\partial x$, it follows from this result that the work done in moving one molecule of solute a distance dx is

$$\frac{\partial n}{\partial x}\,dx\,.\,\frac{kT}{n}$$

so that the force acting on one molecule of solute is $kT/n\ \partial n/\partial x$. We may suppose that the resistance experienced by a molecule moving through the liquid is proportional to the velocity, and we may estimate its order of magnitude by means of Stokes' law, p. 154. This is no more than an "order of magnitude" estimate because the radius of the solute molecule will, at best, be only a few times larger than that of solvent molecules, so that the concept of viscosity is not strictly applicable. We may write

$$\frac{kT}{n}\,\frac{\partial n}{\partial x} = -\,6\pi a\eta v \quad\ldots\ldots\ldots\ldots\ldots\ldots \quad (15)$$

where a is the effective radius of a molecule, η the viscosity, and v the steady velocity of a molecule associated with a concentration gradient $\partial n/\partial x$. (We have the negative sign because the velocity of flow is towards the smaller concentration.) Suppose now that $\partial n/\partial x$ is not constant. Then we consider the transfer of solute across two planes perpendicular to the x-axis, whose abscissae are x and $x + dx$. The number of molecules flowing per second *into* the volume cut off between these planes is $n\,(x)\,v\,(x)$ per unit area across the plane at x, while $n\,(x + dx)\,v\,(x + dx)$ is the number per unit area flowing *out* of the volume across the plane at $x + dx$ per second. Using equation (15) to determine vn at the points x and $x + dx$, the net number of molecules leaving the volume is

$$-\frac{kT}{6\pi a\eta}\left[\left(\frac{\partial n}{\partial x}\right)_{x+dx} - \left(\frac{\partial n}{\partial x}\right)_{x}\right] \text{ per unit area}$$

$$= -\frac{kT}{6\pi a\eta}\left(\frac{\partial^2 n}{\partial x^2}\right)dx \qquad \text{per unit area.}$$

Since the volume cut off between unit area of these two planes is dx, the number of molecules leaving this volume in unit time can also be written $-\,dx\,.\,\partial n/\partial t$, so that we may write finally

$$\frac{\partial n}{\partial t} = \frac{kT}{6\pi a\eta}\,\frac{\partial^2 n}{\partial x^2} \quad\ldots\ldots\ldots\ldots\ldots\ldots \quad (16)$$

If we insert numerical values $k = 1{\cdot}4 \times 10^{-16}$ ergs/degree, $T = 300^\circ$ K, $a \simeq 10^{-8}$ cm., $\eta = 10^{-2}$ poise, we conclude that the constant in equation should be of the order of 10^{-5}, which is of the order found experimentally in liquids. Equation (16) depends basically on the two assumptions, that osmotic pressure is proportional to concentration and that resistance is of "frictional" type, that is proportional to velocity. These would have enabled us to derive equation (16) without knowledge of the constant, but the order of magnitude estimate that we are able to make is of interest. Equation (16) is sometimes known as **Fick's Law,** and a similar equation governs any phenomenon in which a rate of transfer is proportional to a gradient. In particular, it applies to the conduction of heat, and the propagation of electromagnetic waves in metals, and could have been applied to the diffusion of gases, treated in Chapter V (p. 128). Some methods of solving equation (16) are given in the Appendix (p. 298).

Various ingenious methods of measuring the rate of diffusion in liquids have been devised. The most obvious method is to pour the pure liquid carefully on to the top of the solution, and then to follow the gradual penetration of the solute into the pure liquid. This can be done by chemical analysis, but this necessitates the withdrawal of small samples and consequent disturbances. A better plan is to use some optical method, such as following the change in refractive index, or, with a solute such as sugar, the optical rotation, or, with a coloured solution, the growth in the intensity of an absorption line. An ingenious idea, due to **Kelvin,** is to use a number of glass beads of different densities and observe the position of each as the experiment proceeds. This may be regarded as a sampling method, but the disturbances are very small.

Yet another idea has been to use "continuous flow" methods analogous to those used for measurements of heat conductivity and specific heats. If we had a beaker in which pure solvent is kept flowing slowly along the top of a volume of solution, the quantity of solute carried over in unit time would be directly proportional to the diffusivity. Various pieces of apparatus have been devised so that the geometry of the flow is precisely known (the easiest way of doing this is to carry the pure solvent through a tube, contact with the solution being made through a hole in one side), in which case the diffusivity can be calculated numerically.

BIBLIOGRAPHY FOR CHAPTER VIII

R. H. Fowler and E. A. Guggenheim, *Statistical Thermodynamics.* Cambridge University Press (1939).

P. Debye and E. Hückel, *Phys. Zeitschreift,* **24** (1923), 185, 305.

J. K. Roberts, *Heat and Thermodynamics.* Blackie (1951).

J. Frenkel, *Kinetic Theory of Liquids.* O.U.P. (1946).

Examples on Chapter VIII

1. Define what is meant by osmotic pressure. Outline the argument which leads to a connection between the osmotic pressure of a solution and the elevation of the boiling point.

2. What significance do you attach to the fact that one gram-molecule of salt dissolved in a litre of water produces nearly twice as great a depression of the freezing point as does one gram-molecule of sugar? In what circumstances would you expect such a ratio to be exactly two?

3. Calculate the average ionic drift velocity associated with the electrolysis of a solution of sodium chloride containing 1 mol./litre at a current density of 10 amp./cm.^{2}. How would you attempt to detect a difference between the mobilities of the positive and negative ions?

(Avogadro's number $= 6 \times 10^{23}$ mol./g.-mol., charge on the electron $1{\cdot}6 \times 10^{-19}$ coulomb.)

4. Use the method given in the Appendix to derive a solution of the diffusion equation in the form (p. 298) $n = f(t) \sin kx$.

Determine the form of $f(t)$ and show how this solution can be used to estimate the rate at which a difference in concentration between two layers of solution disappears.

5. Give a critical account of the Debye-Hückel theory of strong electrolytes, pointing out the assumptions made and stating whether you consider any of them to be unjustified.

CHAPTER IX

SURFACE TENSION, ADSORPTION, AND ALLIED EFFECTS

1. Introduction

It has long been known that soap films behave in many ways like sheet rubber or other elastic membranes. Thus, (Fig. 78) if we mount a loop of thread on a frame, cover the frame with a soap film and then break the film inside the loop, the remaining film will pull the thread out into a perfect circle (the circle being the curve that has the largest area for a given perimeter), suggesting that the soap film tends to contract as much as possible. It is almost equally easy to show that the surface of pure water or other liquids behaves in a similar way, perhaps the most direct piece of evidence for this being the formation of drops on a tube, which, just before they break away, have the section shown in Fig. 79. On p. 184 we have indicated the theoretical reason for the existence of this strange property of liquid surfaces. This explanation is mainly due to **Thomas Young** and **Laplace**, and, although some of the assumptions they made cannot now be accepted, the basic idea of surface tension being due to intermolecular forces of comparatively short range is so sound physically that Laplace's work can be accepted to-day with appropriate modifications. In this chapter we shall show how surface tension, latent heat, and the van der Waals' constant a, all three of which are manifestations of the intermolecular attractions, can be correlated on the basis of a quite simple theory. We shall also describe a number of methods of measuring surface tension and discuss some problems analogous to those of elastic stability discussed in Chapter IV (p. 103).

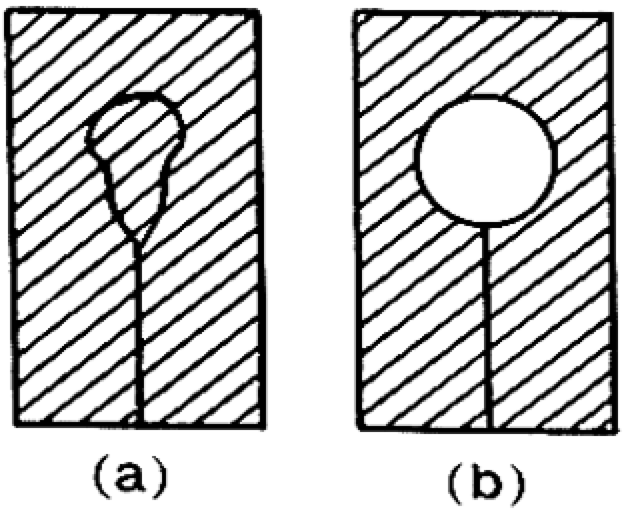


FIG. 78.

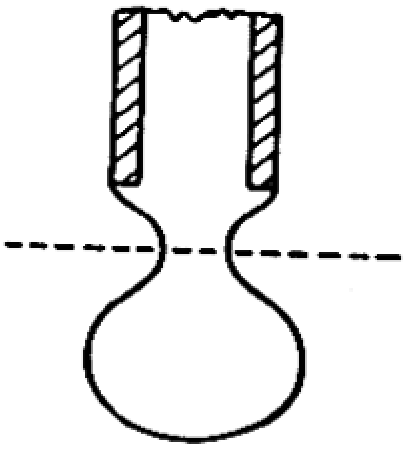


FIG. 79.

2. Surface tension and surface energy. Angles of contact

An experiment such as that illustrated in Fig. 78 suggests that we can associate the contraction of a surface with a decrease of potential energy, so that we are led to think of an *energy per unit surface area.* On the other hand, the existence of the hanging drop shown in Fig. 79 suggests that surface tension is related directly to the *force* that is holding up the drop. If we consider the section of the drop at the height indicated by the dotted line in Fig. 79, the weight of the portion of the drop below this may be thought of as balanced by vertical forces acting tangentially to the surface of the neck.

We are thus led to think of any two elements of a liquid surface as exerting forces on one another across a *line* separating them, so that we arrive at the definition of surface tension as *force per unit length of boundary line.* It might be thought that these two definitions are equivalent, but this is not so. It is very often necessary to allow for a numerical difference between a force that is tending to produce a deformation and the energy change that is associated with the same deformation. Strictly speaking, we ought to do this in discussing moduli of elasticity and elastic energies, but here the difference is not more than a few per cent., except for gases, and is generally neglected. There are, however, other physical situations in which this type of difference must be attended to, for example, the e.m.f. of a reversible electric cell may not correspond with the chemical energy consumed while it is working. For a full discussion of this point, the reader is referred to textbooks on thermodynamics or physical chemistry.

Briefly, the essential physical point is that surface tension decreases with rising temperature. If so, then if a film were stretched adiabatically we can show that a drop in temperature would result, whereas, if we stretch it isothermally it will absorb heat from its surroundings. The change in energy-content of the film is then not precisely equal to the work done. We can investigate this quantitatively by means of the following simple thermodynamic cycle. The processes we invoke can actually be carried out, *e.g.* if we mount a soap-film on a frame with a movable boundary or if we mount such a frame just at the surface of one liquid and form another liquid film inside the frame (p. 256). The cycle we use is as follows:—

(*a*) Stretch the film at a temperature θ. We suppose the increase of energy to be E per unit area. The work done in moving a boundary of length l through a distance x is T_1lx, where T_1 is the surface tension, as defined above, at the temperature θ. The resulting increase in area is lx, so that the mechanical work done is T_1lx, whereas the energy change is E_1lx, so that the heat absorbed must be $(E_1 - T_1)\,lx$.

(*b*) Cool the film at constant area to a temperature θ_2. No work is done, but heat $C(\theta_1 - \theta_2)$ is emitted, where C is the heat capacity of the film.

(*c*) Allow the film to contract at a temperature θ_2 until its area is diminished by lx. The work done by the film is $T_2 lx$, where T_2 is the surface tension at temperature θ_2.

(*d*) Heat the film at constant area to θ_1. No work is done, but heat $C(\theta_1 - \theta_2)$ is absorbed.

Thus, the final result of our cycle is that heat $(E_1 - T_1)\,lx$ has been absorbed at the temperature θ_1, in return for which we have had a net amount of mechanical work $(T_2 - T_1)\,lx$ done by the film. Such a cycle is reversible, and it therefore follows from the general theory of heat engines that

$$\text{Efficiency} = \frac{\text{Work done}}{\text{Heat absorbed at Temp. } \theta_1} = \frac{\theta_1 - \theta_2}{\theta_1} \quad \ldots\ldots \quad (1)$$

i.e.

$$\frac{T_2 - T_1}{E_1 - T_1} = \frac{\theta_1 - \theta_2}{\theta_1}.$$

In the limit, when $\theta_1 - \theta_2$ becomes very small, we have

$$E - T = -\theta\frac{\partial T}{\partial \theta} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (2)$$

A similar equation can be applied to elastic deformations if the moduli vary with temperature. If $\partial T/\partial\theta$ is negative, E is numerically larger than T, which means that a film stretched isothermally absorbs heat. It was for a long time thought that surface tension vanishes at the critical temperature at which the liquid and vapour phases become identical, but **Callendar's** work on the critical behaviour of air-free water suggests that the surface tension really vanishes a little below this. **J. E. Mayer's** theoretical work on the condensation of imperfect gases also makes it very likely that the same thing may occur for other liquids. A similar conclusion can be drawn from the well-known experiment, in which a tube containing liquid carbon dioxide. is gradually heated until the meniscus between liquid and vapour vanishes. (If the surface tension vanished at exactly the critical temperature, we should expect the meniscus to travel to the top or bottom of the tube according to the total mass of carbon dioxide present, unless this corresponded to the exact critical density.)

The angle of contact between liquids and solids.—Clearly we need at least one parameter besides the surface tension to describe the behaviour of a liquid in the neighbourhood of different solids. Evidently, the form of a liquid surface in the neighbourhood of a solid depends on the ratio of the forces exerted on a molecule of liquid by the neighbouring regions of solid and liquid, and three possible situations are shown in Fig. 80. Fig. 80(*a*) shows what occurs when the solid attracts the liquid molecules very strongly—the angle of contact is zero or very small, and we say that the liquid "wets" the solid. This occurs, for example, with very clean glass and water. Fig. 80(*b*) shows a situation sometimes described as "imperfect wetting"—that is to say that the liquid still tries to adhere to the solid, but there is a finite angle of contact. This often occurs with water and a glass surface covered with a trace of grease. The third possible situation is that of an obtuse angle of contact, shown in Fig. 80(*c*) which is usually described by saying that the liquid "does not wet" the solid, the classic example being mercury and glass. Recently it has been shown that certain chemicals, such as soaps, facilitate the wetting of solids by water (mainly by reducing the surface tension of the water and thus reducing the angle of contact). Various numerical measures of "wettability" of solids by liquids have been introduced, but they seem to be practically equivalent to specifying the angle of contact. It sometimes happens that the angle of contact is not a constant, thus it may depend on whether the liquid is advancing over the surface of the solid or receding from it (Fig. 81), and these two angles are sometimes known as the advancing and receding angles of

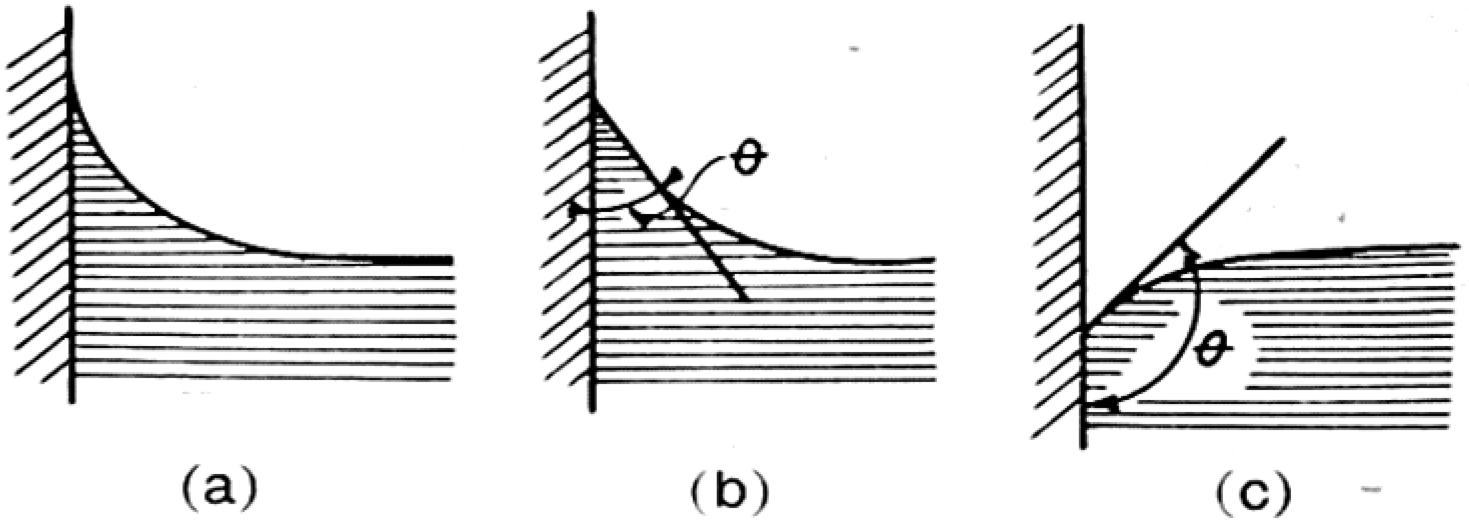


FIG. 80.

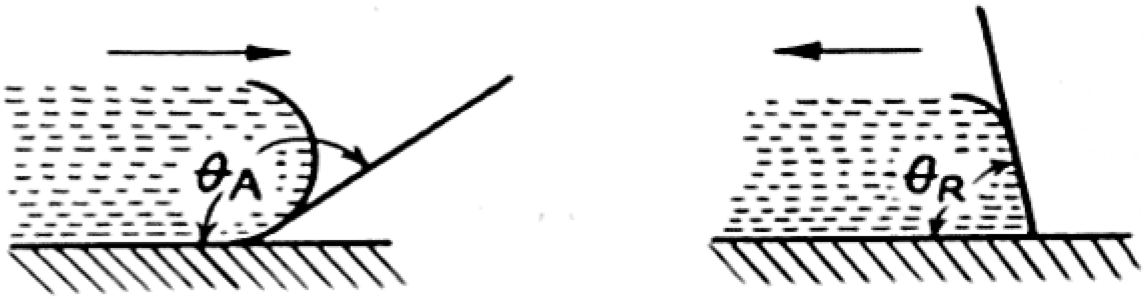


FIG. 81.

contact. This effect is particularly noticeable with mercury on glass, and is sometimes a source of trouble with mercury manometers, which, even when the glass is clean, seem to "stick" and fail to respond to small changes of pressure. For, in a capillary, a change in the angle of contact implies, as a geometrical necessity, a change in the *curvature of the meniscus*, which may correspond to pressures of several millimetres of mercury. We shall now determine the pressure-difference associated with a curved boundary surface.

3. Pressure-differences associated with a curved liquid surface

We shall apply the principle of virtual work (p. 12) to determine the relationship between the pressure-difference and the geometry of a curved surface. Let an element of surface $ABCD$ (Fig. 82) be displaced in such a way that each element moves along the normal to the surface at that point a distance dl. Now we know from three-dimensional coordinate geometry that any sufficiently small element of surface can be described by means of two principal radii of curvature defined in the following way. If the direction of AB is properly chosen in the surface, and BC is at right angles to it, then the normals to the surface at A and C both intersect the normal at B, at O and O' respectively. O is then the centre of curvature of the arc AB and O' that of the arc BC.

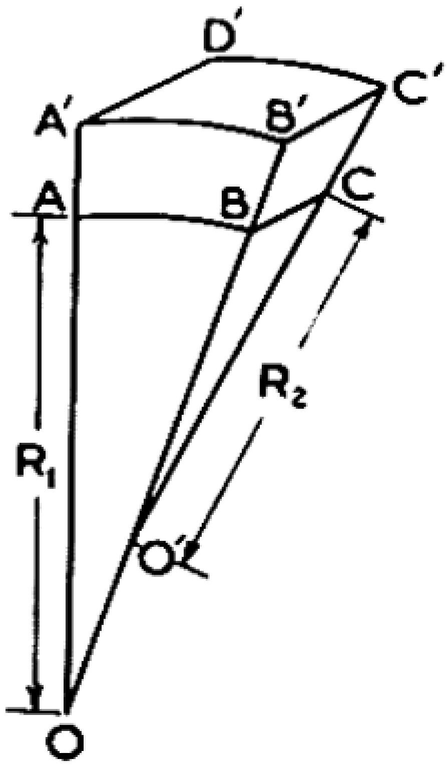


FIG. 82.

If we carry out a virtual displacement of each element of surface by dl along the normal, so that the centres of curvature remain stationary, let $A'B'C'D'$ be the new position of $ABCD$. Work equal to $(p - p')\,AB.CD.dl$ is done by the pressure-difference on the two sides of the surface as a result of its moving a distance dl normal to itself, $AB.CD$ being the approximate area of the element and p, p' the pressures on its two sides. At the same time the area of the surface has changed because $A'B'$ is longer than AB, and, as Fig. 82 is drawn, $C'D'$ is longer than CD. If $OB = R_1$ we have

$$\frac{A'B'}{AB} = \frac{R_1 + dl}{R_1}$$

since O is the centre of curvature of both arcs. Similarly, if $O'B = R_2$, we have

$$\frac{C'D'}{CD} = \frac{R_2 + dl}{R_2}.$$

The change of area is approximately

$$A'B'.C'D' - AB.CD = AB.CD\left[\left(1+\frac{dl}{R_1}\right)\left(1+\frac{dl}{R_2}\right)-1\right]$$

$$= AB.CD\,dl\left(\frac{1}{R_1}+\frac{1}{R_2}\right)+O\,(dl)^2.$$

Work equal to $T \times$ change in area is done against the mechanical force of surface tension (we must suppose the change to take place adiabatically), and, by the principle of virtual work, the total mechanical work done during this imagined displacement must vanish to the first order in dl, that is

$$(p-p').AB.CD.dl - T.AB.CD.dl\left(\frac{1}{R_1}+\frac{1}{R_2}\right)=0,$$

i.e.

$$p - p' = T\left(\frac{1}{R_1}+\frac{1}{R_2}\right) \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (3)$$

Equation (3) applies if R_1 and R_2 are of the same sign, that is to say if O and O' are on the same side of the element $ABCD$, and the sign of equation (3) shows us that we get the greater pressure on this side of the surface. Such surfaces, of which the sphere and ellipsoid are typical, are called *synclastic*. A second type of surface formed by soap films, called *anticlastic*, is a surface like a saddle in which the principal centres of curvature are on *opposite sides* of the element. The investigation can be repeated for such a surface, and we can show that equation (3) still holds provided that we consider R_1 and R_2 to be of opposite signs. If they are numerically equal, then there is no change in pressure across such a surface, otherwise the pressure change will be of a sign determined by the *numerically smaller* of R_1, R_2. Subject to a suitable sign convention, equation (3) therefore holds universally.

Equation (3) holds at any liquid-gas or liquid-liquid boundary, the appropriate value of T being used in each case. In a soap film, $p - p'$ represents the difference between the pressures in the gas and liquid, so that if, as is usually the case, the two surfaces of the film are nearly parallel, the pressure-difference across the whole film will be double that given by equation (3). We now consider some special forms of surface.

(*a*) *Spherical bubble or drop.*—For each surface $R_1 = R_2 = R$, the radius of the bubble, and equation (3) gives us at once for the excess pressure inside the bubble the result

$$p_1 - p_2 = \frac{2T}{R}+\frac{2T}{R}=\frac{4T}{R} \quad \ldots\ldots\ldots\ldots\ldots \quad (4)$$

This result can be obtained, without resort to the general theory, by a very simple argument (Fig. 83). Consider the equilibrium of the right-hand half of the bubble, divided by a vertical diameter plane. The section of the bubble is a circle of radius R, and the total force holding the two halves of the bubble together is $2T \times 2\pi R$ ($2T$ because there are two surfaces). This must be balanced by the force associated with the excess pressure inside the bubble. The force on an element whose polar coordinates are (R, θ) acts along a radius, that is, its horizontal component is $(p_2 - p_1) \cos \theta dA$, and the area of a ring-shaped element cut off between θ and $\theta + d\theta$ is $2\pi . Rd\theta . R \sin \theta$, so that the total horizontal force on the right-hand half is

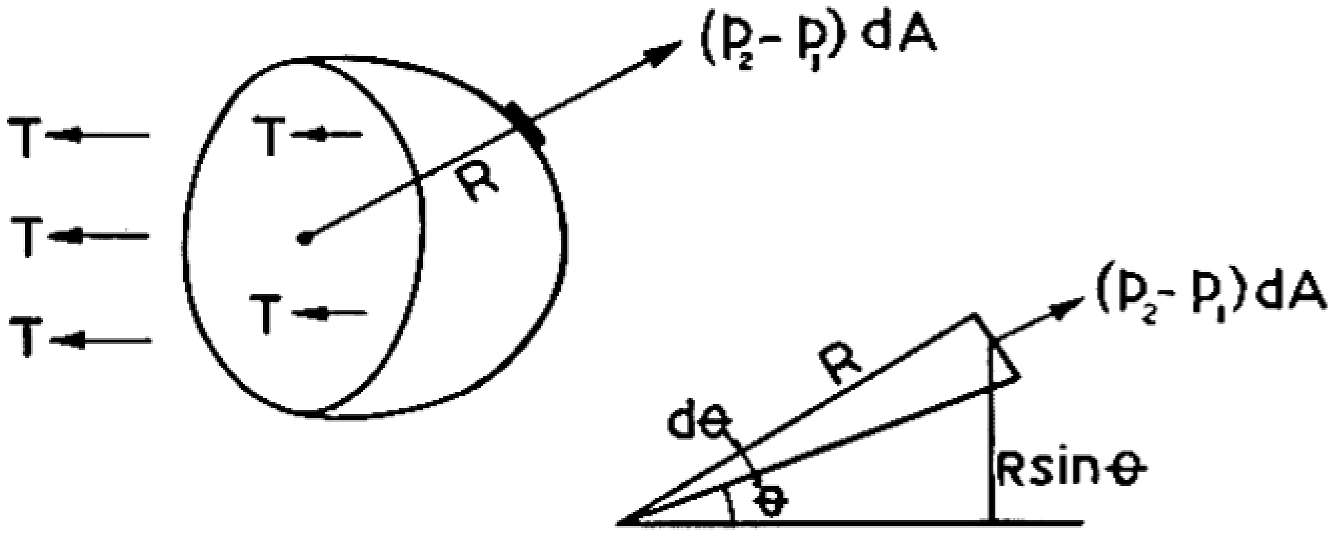


FIG. 83.

$$(p_2 - p_1)\int_0^{\pi/2} \cos \theta . 2\pi R d\theta . R \sin \theta = (p_1 - p_2)\, \pi R^2,$$

which has to be equated to the surface tension force $4\pi RT$, giving us equation (4). For the excess pressure inside a drop of liquid, we should halve the above value because there is only one boundary surface

$$p_2 - p_1 = \frac{2T}{R} \quad \ldots\ldots \quad (5)$$

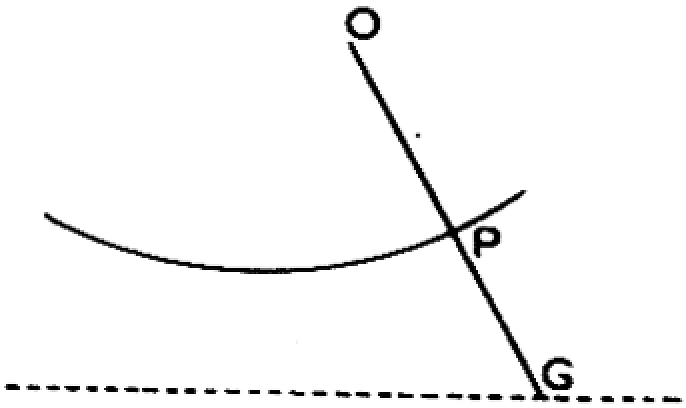


FIG. 84.

(*b*) *Catenoid soap films.*—Such a film can be formed, *e.g.* between the mouths of two coaxial funnels (Fig. 84). If both sides of the film are open to the atmosphere, there can be no pressure-difference across it, so that the two principal radii of curvature are equal and opposite. The two centres of curvature are at O and G (Fig. 84), where G is the point where the normal at P cuts the axis of the surface of revolution, and O is the centre of curvature of a plane section of the curve.

Thus we must have $OP = PG$, that is, if y is expressed as a function of x,

$$\frac{\left[1+\left(\frac{dy}{dx}\right)^2\right]^{\frac{3}{2}}}{\frac{d^2y}{dx^2}} = y\left[1+\left(\frac{dy}{dx}\right)^2\right]^{\frac{1}{2}} \quad \ldots\ldots\ldots\ldots (6)$$

using the known expressions for the radius of curvature, and the intercept of the normal on the x-axis. This gives us—

$$y\frac{d^2y}{dx^2} = 1+\left(\frac{dy}{dx}\right)^2$$

or

$$\frac{\frac{dy}{dx}\frac{d^2y}{dx^2}}{1+\left(\frac{dy}{dx}\right)^2} = \frac{1}{y}\frac{dy}{dx}.$$

Integrating, $\frac{1}{2}\log\left[1+\left(\frac{dy}{dx}\right)^2\right] = \log\frac{y}{a}$ where a is a constant

or

$$1+\left(\frac{dy}{dx}\right)^2 = Ay^2.$$

The general solution of which is

$$y = b\cosh\left(\frac{x+c}{b}\right) \quad \ldots\ldots\ldots\ldots\ldots\ldots (7)$$

where b and c are constants. Equation (7) is the equation of a *catenary*, that is the curve in which a heavy chain hangs when suspended between two points. Such films connecting two funnels are usually known as catenoid films. If the radii are equal, we may, by symmetry, take the origin half-way between the funnels, and suppose that c is their distance apart. Then b is given by the equation $R = b\cosh(c/2b)$, where R is the radius of one funnel. This equation does not have a solution for all values of R and c, the minimum possible value of R for a given value of c being given by

$$R = \frac{c}{2q}\cosh q$$

where q is the solution of the equation $x\tanh x = 1$. This result means that *a catenoid film cannot be formed between two funnels that are more than a certain distance apart*, so that such a film would certainly break if we drew them apart beyond this distance, however careful we were to avoid disturbances. This behaviour is similar to that of *elastic instability* which we discussed in Chapter IV, p. 103. We shall meet similar cases below, pp. 239, 240.

(c) *Nearly cylindrical films.*—The catenoid film is almost the only case that can be worked out without approximation. If there were a finite pressure-difference across the film, equation (3) gives us

$$2T\left(\frac{1}{OP}-\frac{1}{PG}\right)=p_1-p_2=\text{constant}$$

and this becomes very much more difficult to integrate. However, if we suppose the pressure adjusted so that the film is nearly cylindrical, we can make the same approximation that we made in the theory of elasticity (p. 93), namely, we suppose that dy/dx can always be neglected compared with unity. In this case, we have

$$\frac{d^2y}{dx^2}-\frac{1}{y}=\frac{p_1-p_2}{2T}.$$

We now make use of the fact that the film is nearly cylindrical, so that we may put $y=R+z$, where z is a small quantity whose square we neglect. Our equation then becomes, neglecting z^2,

$$\frac{d^2z}{dx^2}+\frac{z}{R^2}=\frac{1}{R}+\frac{p_1-p_2}{2T} \quad \ldots\ldots\ldots\ldots\ldots \quad (8)$$

This equation is of the same form as the equation for the simple harmonic oscillator treated in the Appendix (p. 285). The complementary function is

$$z=A\cos\left(\frac{x}{R}\right)+B\sin\left(\frac{x}{R}\right)$$

where A and B are constants. A particular solution is obviously $z=R+\frac{(p_1-p_2)R^2}{2T}$, so that the complete solution of (8) is

$$z=R+\frac{(p_1-p_2)R^2}{2T}+A\cos\left(\frac{x}{R}\right)+B\sin\left(\frac{x}{R}\right) \quad \ldots\ldots \quad (9)$$

Now the constants A and B have to be determined so that $y=R$, that is $z=0$, at the two ends of the film, which we may suppose to be at $x=0$ and $x=l$. We get from these two conditions

$$A=-R-\frac{(p_1-p_2)R^2}{2T},$$

$$B=-\left(\frac{1-\cos(l/R)}{\sin(l/R)}\right)\left(R+\frac{(p_1-p_2)R^2}{2T}\right) \quad \ldots\ldots\ldots \quad (10)$$

Equations (10) imply that $B/A=\tan(l/2R)$ and it can easily be verified that this implies $dz/dx=0$ at $x=\frac{1}{2}l$, that is to say that z has a maximum or minimum there.

Differentiating equation (9) twice

$$\frac{d^2z}{dx^2} = -\frac{A}{R^2}\cos\left(\frac{x}{R}\right) - \frac{B}{R^2}\sin\left(\frac{x}{R}\right).$$

Putting $x = \frac{1}{2}l$ and substituting for A and B from (10) we get

$$\left(\frac{d^2z}{dx^2}\right)_{x=\frac{1}{2}l} = -\left(\frac{p_2 - p_1}{2T} - \frac{1}{R}\right)\left[\sec\left(\frac{l}{2R}\right)\right] \quad\dots\dots\dots (11)$$

the second bracket being positive or negative according as $\cos l/2R$ is positive or negative, *i.e.* according as $l < \pi R$ or $> \pi R$. The solutions for $l > 2\pi R$ would imply a film that had several maxima and minima of z along its length, and can be ruled out on physical grounds. Similarly, in problems of elastic stability, p. 105, and whirling of shafts, p. 106, a solution which implies *more than one* turning point is not usually physically significant, because failure would have already occurred under less onerous conditions.

Equation (11) enables one to deduce the following results:—

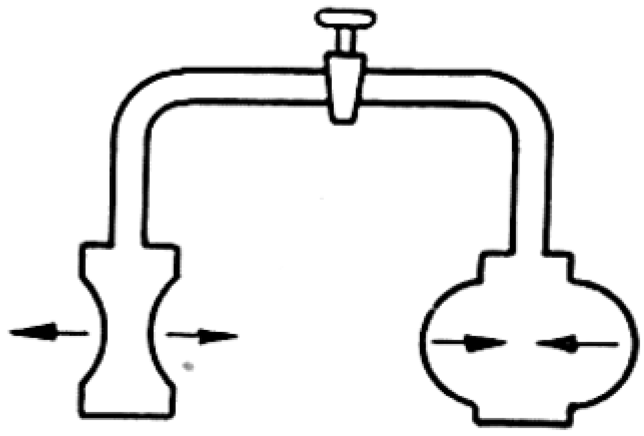


FIG. 85.

(i) $l < \pi R$. d^2z/dx^2 positive, that is z a *minimum* at $x = l/2$, implies $p_2 - p_1 < 2T/R$. An *inwardly bulging* film has a smaller excess of pressure inside than the $2T/R$ corresponding to a cylindrical film. Similarly, an outwardly bulging film has a larger excess of pressure than the cylindrical film.

(ii) $l > \pi R$. The above results are just reversed because of the change in sign of $\cos(l/2R)$. An outwardly bulging film has a smaller excess of pressure than a cylindrical film, which has in turn a smaller excess than an inwardly bulging film.

From (i) and (ii) it is possible to deduce that a cylindrical film *longer than its circumference* would be unstable, for a small disturbance could convert it into the equivalent of a "bulged-in" and "bulged-out" film placed "end to end", each longer than half the circumference, and (ii) shows that the resulting pressure changes would tend to increase the bulgings. On the other hand, a film shorter than its circumference would by (i), be stable against such a disturbance. These conclusions were all experimentally verified by **C. V. Boys.** If two films of the same length, shorter than the half-circumference, are put into communication with one another, air passes to a concave film from one that is less concave, straight, or convex (Fig. 85), while, if the films are longer than half the circumference, a concave film will *close up*, blowing air into a straight or convex one.

A similar experiment carried out with spherical bubbles leads always to the smaller one closing up and expanding the large one, because the excess pressure in a spherical bubble is inversely proportional to the radius (p. 237).

The results on cylindrical films have an important bearing on the stability of jets, and it is also almost certain that the critical radius of a drop (Fig. 79) is settled by the point at which the neck becomes unstable. One way, not however by any means the only one, in which a jet breaks up is shown in Fig. 86, transverse or "varicose" oscillations increasing in amplitude along the length until the jet breaks up completely, into large drops of about the same diameter as the jet, and an almost equal number of smaller ones, made up of the remains of the "necks". A similar small drop, "Plateau's spherule", is often observed in the situation shown in Fig. 79. In the jet, the effect of viscosity opposes that of surface tension, and it is only fluids of high viscosity that can form long thin filaments.

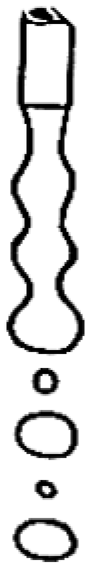


FIG. 86.

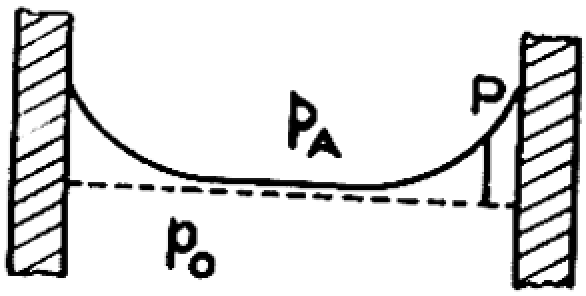


FIG. 87.

(*d*) *The form of the free surface of a liquid.*—We consider the simplest case, a liquid in equilibrium under gravity between two vertical solid surfaces. Now the pressure in the liquid itself will be the same, p_0 say, all along the dotted line, which we take as the x-axis. The pressure p_A in the atmosphere is constant. The pressure at the point P just inside the liquid is (Fig. 87)

$$p_A - 2T\frac{d^2y}{dx^2}\left[1+\left(\frac{dy}{dx}\right)^2\right]^{-\frac{3}{2}},$$

and it is also equal, by hydrostatics, to $p_0 - \rho g y$, so that the equation of the free surface is determined by

$$\frac{d^2y}{dx^2}\left[1+\left(\frac{dy}{dx}\right)^2\right]^{-\frac{3}{2}} - \frac{\rho g y}{2T} = \text{constant} \quad \ldots\ldots\ldots \quad (12)$$

We can integrate this once by multiplying through by dy/dx. The result is

$$\left[1+\left(\frac{dy}{dx}\right)^2\right]^{-\frac{1}{2}}-\frac{\rho g y^2}{2T}=Ay+B \quad \ldots\ldots\ldots\ldots \quad (13)$$

(The expression in the bracket is $-\cos\psi$, where ψ is the angle the tangent to the curve makes with the x-axis.) This could be integrated again in terms of elliptic integrals, but we can use an approximation that is suitable for regions where the surface is nearly flat. We consider only the first two terms of the expression in the bracket, neglecting $(dy/dx)^4$.

On this basis equation (13) may be written

$$\left(\frac{dy}{dx}\right)^2=\frac{\rho g}{T}\left((y-a)^2-b^2\right) \quad \ldots\ldots\ldots\ldots \quad (14)$$

where a and b are new arbitrary constants. The solution of this is

$$\frac{y-a}{b}=\cosh\left\{\left(\frac{\rho g}{T}\right)^{\frac{1}{2}}(x-c)\right\} \quad \ldots\ldots\ldots\ldots \quad (15)$$

The constants a and c depend on the choice of origin in relation to the surface, but the constant b is more difficult. For a liquid with a finite angle of contact with the wall, b could be chosen so that dy/dx had the corresponding value at the solid surface, but approximation (14) becomes useless in the neighbourhood of a solid whose angle of contact is zero or small because there dy/dx is *large*. We can determine the form of the surface in the immediate neighbourhood of the solid, by making the *opposite* approximation to (14), that is we neglect 1 in comparison with dy/dx. This gives

$$-\frac{dx}{dy}=\frac{\rho g y^2}{2T}+Ey+F,$$

or

$$x=G-\frac{Ey^2}{2}-F-\frac{\rho g y^3}{6T} \quad \ldots\ldots\ldots\ldots \quad (16)$$

The constant G depends on the choice of origin, the constant F on the angle of contact (which determines dx/dy at the solid), while the constant E has to be determined indirectly from the conditions at the other boundary. One can get a sufficiently good answer in practice by using (16) in the neighbourhood of the solid boundaries, (15) in the middle of the surface, and then "fitting together" these two solutions choosing the constants, E and b which are not determined by the choice of origin, and so are at our disposal, so that, at some intermediate point, y and dy/dx are both continuous. (This process of "fitting together" the solutions corresponding to two limiting cases is typical of mathematical physics.)

It will be seen, even in this simple case, that the calculation of the form of the meniscus is quite involved, and the calculation of the form of the meniscus in a cylindrical tube is even more so, although **Bashforth** and **Adams** have made tables which enable the form of axially symmetrical menisci to be computed numerically in many cases. The assumption usually made is that, for a zero angle of contact the shape can be regarded as a hemisphere, or, for a finite angle of contact, the appropriate spherical cap. In very accurate work, data of Bashforth and Adams can be used to improve upon these assumptions. For a zero angle of contact (Fig. 88), the rise to be expected can be calculated quite simply. If R is the tube, the upward force is $2\pi RT \cos\theta$, where θ is the angle of contact, and this is balanced by the weight of a column of water of height h, where h is measured to the bottom of the meniscus as usual, together with the weight of water above the level of the bottom of the meniscus. The volume of this, with the "hemispherical" assumption and zero angle of contact, is half the volume of the hemisphere, that is $\pi R^3/3$, so that we have, if $\theta = 0$,

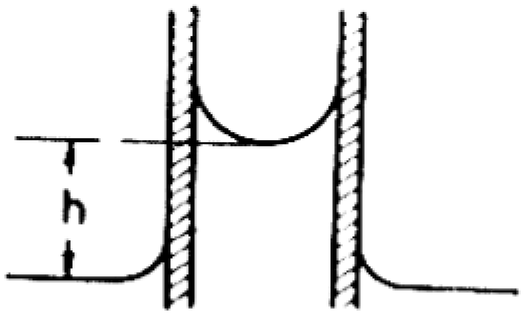


FIG. 88.

$$2\pi RT = \left(\pi R^2 h + \frac{\pi R^3}{3}\right)\rho g \quad \ldots\ldots\ldots\ldots \quad (17)$$

so that the rise in the tube is known in terms of T. The mass of liquid above the bottom of the meniscus can be estimated by an analogous geometrical argument if θ is finite. **Lord Rayleigh** gave more accurate formulae for use when R is not small compared with h.

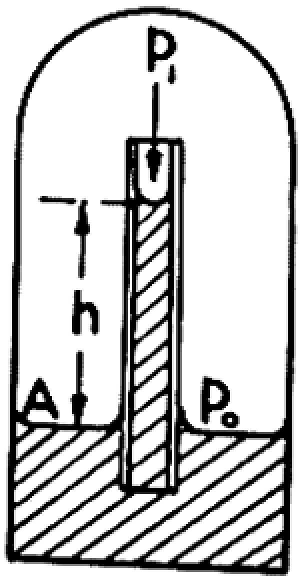


FIG. 89.

(*e*) *Change of vapour pressure due to curvature of surface.*—So far we have considered the most obvious property of surface tension which is that of determining the form of a liquid surface. If, however, it is really a manifestation of intermolecular attractions one might expect the departure of a molecule from the surface to be helped or hindered according to its curvature. It is possible to deduce a relationship between surface tension and vapour pressures, independent of any particular theory of liquids, quite analogous to that deduced between osmotic pressure and vapour pressure (p. 211). To do this, we imagine the capillary rise experiment done under a bell-jar, so that only vapour is present (Fig. 89). Then, if p_0 is the vapour pressure over the flat surface, and p_1 that over the curved surface, we may imagine a molecule leaving the liquid at the top of the tube, transported to the bottom of the tube, entering the liquid through the flat surface, and then being carried up the tube to the top again, the net work done during this cycle being zero. Put in another way, we can say that the pressure on the liquid side of the meniscus is $p_0 - \rho gh$, where ρ is the density of liquid, and the term ρgh allows for the height of the column of liquid. The pressure on the vapour side of the meniscus is $p_0 - \sigma gh = p_1$, where σ is the density of vapour. Since the liquid surface at A (Fig. 89) is flat, the pressure is the same on both sides of it. Thus the difference in pressure on the two sides of the meniscus is

$$(\rho - \sigma)\, gh = \frac{2T}{r} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (18)$$

But

$$p_1 = p_0 - \sigma gh = p - \frac{\sigma}{\rho - \sigma}\frac{2T}{r} \quad \ldots\ldots\ldots\ldots \quad (19)$$

As with osmotic pressure, this result should be corrected for the compressibilities of the gas (and liquid) when very long columns are involved.

This formula was due to **Kelvin** and has been applied to a discussion of the effect of surface tension on the evaporation of drops. It is not altogether clear that this is justified, though the result is probably qualitatively correct. Apart from the difficulty that, for the effect on vapour pressure to become appreciable, the drop must be 10^{-2} cm. or less in diameter, and the effects of compressibility of the vapour and liquid then become important, it is by no means obvious that these considerations apply to a *convex* liquid surface when they have been derived for a *concave* one. The Laplace theory of surface tension does make this extremely plausible, but it does seem clear that some such justification is required. Very recently, a direct check of the Kelvin formula has been made by La Mer and Gruen, the method being to balance the *increase* of vapour pressure due to the curvature of small drops against the *decrease* of vapour pressure due to the presence of a dissolved solid. If a drop of solution is allowed to come into equilibrium with pure solvent in bulk, via the vapour phase, it is not difficult to show that the equilibrium radius is uniquely determined by the quantity and nature of solute present in the drop. (Since the vapour is pure solvent, a change of radius of the drop implies a change of concentration.) The reader is referred to the original paper for details.

A simple argument, that gives much the same result as Kelvin's treatment, does exist. The energy associated with the surface of a drop of radius R is evidently $S.4\pi R^2$, while the energy needed to evaporate the drop is $L.4\pi R^3/3$, where L is the latent heat per unit volume. Thus a *small enough drop would evaporate spontaneously*, the surface energy being sufficient for this. The critical radius is evidently $3S/L$, where L is to be expressed in ergs/cm.^{3}, so that for water

$$R \simeq \frac{3 \times 100}{600 \times 4{\cdot}2 \times 10^7} \simeq 1{\cdot}2 \times 10^{-8} \text{ cm.} \quad \ldots\ldots\ldots \quad (20)$$

which gives us an estimate of this critical size of drop. As will be seen, R is comparable with an atomic diameter, which is physically reasonable, and it also implies that, even if we have a considerably supersaturated vapour, a drop containing several molecules must be formed before condensation can start. This is in line with the experimental facts about nuclei summarised in Chapter VII, p. 193.

Very similar considerations can be applied to the growth of gas or vapour bubbles in a liquid supersaturated with dissolved gas, or in a liquid under tension. For, suppose we have a gas-bubble of radius 10^{-6} cm. The pressure-difference across the surface is $2T/R \sim 2 \times 10^8$ dynes/cm.^{2} or 200 atmospheres. Thus a bubble of this size would collapse unless, *e.g.*, there was a pressure in the bubble, or a tension in the liquid, of at least this. On the other hand, if such a pressure-difference did exist, the bubble would expand rapidly. In this way one can get a qualitative understanding, from a different angle, of the tensile strength and possibility of superheating of a liquid, already discussed on p. 192, and of the related phenomenon of the formation of supersaturated solutions of gas in liquid. It is well known that soda-water can be kept for a long time in a smooth glass, which offers no nuclei on which bubbles can start, and it has been shown that the pressure over a solution of gas in liquids, formed at a pressure of 100 atmospheres, can, with precautions, be reduced to 1 atmosphere without effervescence occurring. Attempts have been made, with fair success, to estimate the size of bubble that might be expected to form spontaneously as the result of random fluctuations, and the resulting predictions are in fair agreement with the facts.

(*f*) *The forces between small floating bodies due to surface tension.*—The observation that small bubbles on the surface of a liquid, or small bodies such as corks and straws, usually attract one another, and travel towards the sides of the vessel, is an elementary one. The reason for this behaviour can be understood if we consider, for simplicity, two objects with vertical sides. Three cases are possible: the liquid "wets" both [Fig. 90(*a*)], the liquid "wets" neither, *i.e.* has an obtuse angle of contact [Fig. 90(*b*)], the liquid wets one and not the other [Fig. 90(*c*)]. In Figs. 90(*a*) and (*b*) we shall show that the portion of liquid above the bottom of the meniscus can be ignored for the purpose of calculating horizontal forces. (The argument is similar to the familiar argument in hydrostatics, which shows, by considering the equilibrium of a prism-shaped volume of liquid, that the pressure in a liquid is the same in all directions.)

Consider the equilibrium of the portion of liquid whose section is ABC, Fig. 90(a). The horizontal forces acting on a portion of length l are $-Tl$, due to the surface tension acting along the line through A perpendicular to the paper, the pressure of the atmosphere, which acts over the curved surface whose section is AB, the sum of the horizontal components of the force due to this is simply $p_a . l . BC$, and these are to be together equal to the horizontal component of the force exerted by the left-hand body on this portion of liquid. This force is then $(p_a . BC - T)\, l$. By Newton's third law this is also equal to the force exerted by the portion of liquid on the left-hand body, but this is just what we should calculate for the force acting on the body if we supposed the liquid *continued horizontally right up to the surface of the body* as shown in Fig. 90(a), the term Tl representing the effect of surface tension acting horizontally, while the term $p_a . BC . l$ is just the extra force that would result from the pressure of the atmosphere on the area $BC . l$ which would be exposed by the removal of the meniscus. We can apply the same argument to the meniscus between the two bodies, and we notice that the argument still holds even if the angle of contact is finite. An exactly similar argument applies in Fig. 90(b). In both these figures the terms Tl have no resultant effect, because they are the same over the left and right of the bodies. The hydrostatic pressures differ because C and D are at different heights on the two sides. The pressures at C and D would be atmospheric if the menisci were suppressed, and the following considerations will be on this basis. In both figures, let C' be the point at the same level as C.

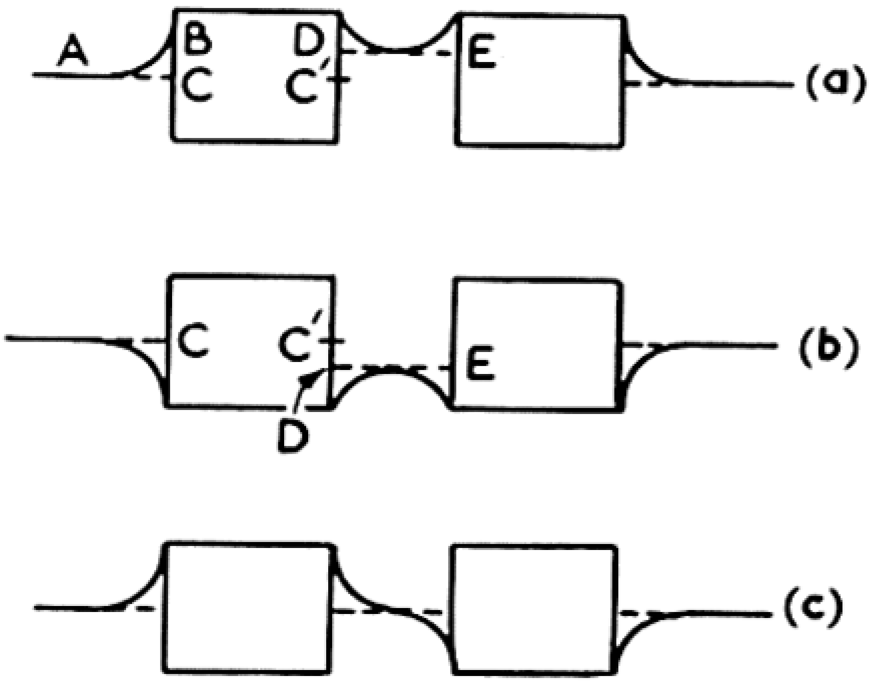


FIG. 90.

In Fig. 90(a) the hydrostatic pressures are equal on both sides of the left-hand body everywhere *below* the line CC', but *above* C we have the atmospheric pressure, while above C' we have a hydrostatic pressure that is less than the atmosphere until we reach D, where it becomes equal to atmospheric. Thus the bodies are forced together.

In Fig. 90(b), we have the atmospheric pressure on both sides of the left-hand body everywhere *above* the line CC', but, between C' and D

the pressure is atmospheric as it is at C, whereas below C the hydrostatic pressure becomes greater than the atmospheric. Thus again the bodies are forced together.

In Fig. 90(*c*), we can suppress the menisci on the outside edges of the two bodies and still calculate the horizontal forces correctly, but we cannot do this for the inside menisci because the surface of the water is never horizontal. We *can*, however, modify the argument already given, and show by the same method that we can calculate the correct forces if we replace the actual surface not by the dotted line in Fig. 90(*c*), but by the *sloping* tangent to this surface. There is now little difference in the hydrostatic pressures on the two sides of the left-hand body, but there is a difference in the surface tension forces, because the force on the left of the left-hand body is acting horizontally, while that on the right of the left-hand body acts at an angle to the horizontal. Thus the forces are not in equilibrium, and the bodies are forced apart.

These considerations can be verified in a beautiful way by a simple experiment. We may float balls of cork and paraffin wax on the surface of water and the sides of the vessel attract the cork balls and repel the paraffin wax ones. If now we "crown the bowl", *i.e.* fill it so full that the upper surface becomes convex instead of concave, we shall produce the same form of surface near the walls as we should get if the water did not wet them. The cork balls then leave for the middle of the vessel, while the paraffin balls are attracted to the sides.

4. Methods of measuring surface tension

We may distinguish a number of different cases:

(*a*) The film, *e.g.* of soap solution, that can be studied in the absence of bulk liquid.

(*b*) The surface of a pure liquid.

(*c*) The surface of a contaminated liquid.

(*d*) The interface between two liquids in bulk.

In case (*a*) methods are available that are not possible for a liquid in bulk. We may measure, for example, the excess pressure inside a soap bubble, whose radius is simultaneously measured, *e.g.* by an optical method using its surface as a curved mirror. The pressure excess is of the order of a millimetre head of water, but this is easily measurable with a manometer and travelling microscope (or the movement of the meniscus in the manometer can be increased by the device of making one arm nearly horizontal).

In addition, we may be interested in the variation of surface tension with thickness. This may be investigated by an apparatus similar to that shown in Fig. 85, in which we begin with two nearly straight films, and allows one to drain naturally but keep the other one thick (which can be done by passing an electric current up it). (This is an instance of the phenomenon of *cataphoresis*, or the relative motion of liquid and suspended particles that occurs if an electric field is applied to colloidal solutions. The reader is referred to textbooks of physical chemistry for a discussion of the phenomenon, which may be traced to the fact that colloidal particles are usually electrically charged.) The thickness of the film that is allowed to drain can be followed by optical interference methods, and any variation in the surface tension is revealed by a change in pressure inside the film, and a consequent change of shape of the nearly cylindrical films.

In practice, it is found that, as long as the film has a thickness comparable with the wavelength of visible light ($\sim 5 \times 10^{-5}$ cm.) the change in surface tension is extremely small, the experiment described above enabling an upper limit of $\frac{1}{2}$ per cent. to be put on it. If the film is allowed to drain further, "black spots" suddenly appear on it, and finally the whole film becomes black by reflected light.

Once the film becomes thinner than a quarter of the wavelength of light, it is no longer capable of showing interference colours and it is necessary to resort to other methods of measuring the thickness. This has been done by various methods, for example, by measuring the electrical resistance of a black film or by forming a large number of such films in a tube, and measuring the combined optical thickness of them all by a Jamin interferometer. (The reader is referred to textbooks of light for a description of this instrument.) The thickness of a black film is generally accepted to be in the neighbourhood of $1{\cdot}2 \times 10^{-6}$ cm. with some evidence that a film of half this thickness is also possible. These figures are consistent with the black film being either one or two molecules thick, the soap molecules lying with their lengths perpendicular to the film. A fatty acid consists of a long hydrocarbon chain, containing 15-20 CH_2 groups and ending in a COOH group, while, in a soap, the hydrogen in the OH group is replaced by a metal. Since the distance between carbon atoms in a hydrocarbon chain is known, from X-ray evidence, to be between 3 and 4×10^{-8} cm., it will be seen that the observed thickness of a black film is consistent with its soap molecules being arranged lengthwise in this way.

The surface tension does not diminish steadily as the thickness of the film decreases, instead the behaviour is probably that shown schematically in Fig. 91, in which surface tension T is plotted against thickness of film t. For thicknesses above those corresponding to the point P, the tension hardly changes with thickness. The interpretation of this is that, for a sufficiently thick film, the two surfaces act independently of one another. As the thickness diminishes, there must come a time when this ceases to be true and there is probably a range of thicknesses for which pulling out the film weakens it (the region *PQ* in Fig. 91). Such a region cannot, however, be experimentally realised, just as regions such as *FG* and *LM* (Fig. 74, p. 192) cannot be realised in a condensing van der Waals' gas.

If this region where dT/dt was positive continued right down to the origin, there would evidently be no possibility of a stable film of thickness less than that corresponding to *P*, because any portion of film falling below this thickness would be dragged out indefinitely by the larger constant tension in the surrounding thicker film. It is indeed an observed fact that films of a thickness between that of the black film ($\sim 10^{-6}$ cm.) and that of the coloured films (about 10^{-5} cm. and upwards) cannot be realised experimentally, and, moreover, that in a coloured film with "black spots" there is a sharp dividing line at the edges of the black spots, not a gradual change over intermediate thicknesses. These facts can be explained if we suppose that the variation of surface tension with thickness is as shown in Fig. 91, the portion *QRS* representing a stable region corresponding to the black film. The point *R* corresponds to the thickness of black film that can be in equilibrium with the thicker coloured films, because the tensions are the same, whereas the range of thicknesses between *Q* and *S* would correspond to possible stable thicknesses of a film that was entirely black. The ranges *OS* and *QP* would be unstable.

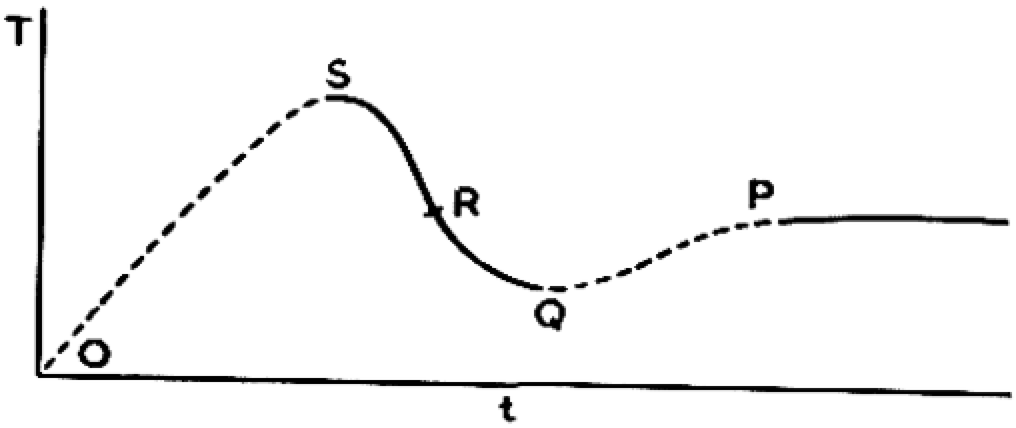


FIG. 91.

This assumed trend of curve certainly explains the existence of black films. We may compare the proposed model with the behaviour of a van der Waals' liquid already referred to (p. 192) in which the portion of the curve for which $(\partial P/\partial V)_T$ is positive, cannot be realised experimentally. We may also compare it with the behaviour of a steel specimen when stressed beyond its yield-point, where the stress-strain curve becomes practically horizontal for a time, but is succeeded by a further rising portion, so that the specimen does not break actually at its yield-point, but can be extended considerably further (p. 76).

Such a model is not, however, very satisfying unless we can suggest some plausible explanation for such peculiar behaviour of T as a function of t. This can be done. Studies of unimolecular films (see p. 256) of fatty acids on the surface of water do suggest that the long paraffin chains cannot enter the water structure without thoroughly disturbing it, and breaking up a great many of the hydrogen bonds (p. 207) that link neighbouring water molecules together. We know that paraffins themselves (hydrocarbon chains without a COOH group) are almost completely insoluble in water, which, as we have seen from the theory of solutions (p. 219) implies that work would have to be done to mix up these two types of molecule. However, the COOH group is capable of forming hydrogen bonds with water molecules, and indeed, the lower fatty acids, and their salts with the alkali metals, have only a short hydrocarbon chain and dissolve in water freely. The picture that one forms of a thin soap film is then shown in Fig. 92, the hydrocarbon chains "standing out" from the water surface, while the COONa groups, supposing that we have a sodium soap, are interspersed with water molecules. Now, if the film is thick, the two sides will have no influence on one another. This will correspond to the region to the right of P in Fig. 91. As the film becomes thinner, the COONa groups on the two sides will come within one another's influence, and, if there is a strong *attraction* between them, one can understand the diminution of work necessary to stretch the film. This stage corresponds to the region PQ in Fig. 91. When the COONa groups come nearly into contact, the film cannot be extended further without pulling the soap molecules apart sideways, and this involves either separating the long hydrocarbon chains, or else separating neighbouring COONa groups or both, and this will require work, and in this way it is possible to understand the existence of the portion QRS of the curve. This corresponds to the most usual type of "black" film, but, in some

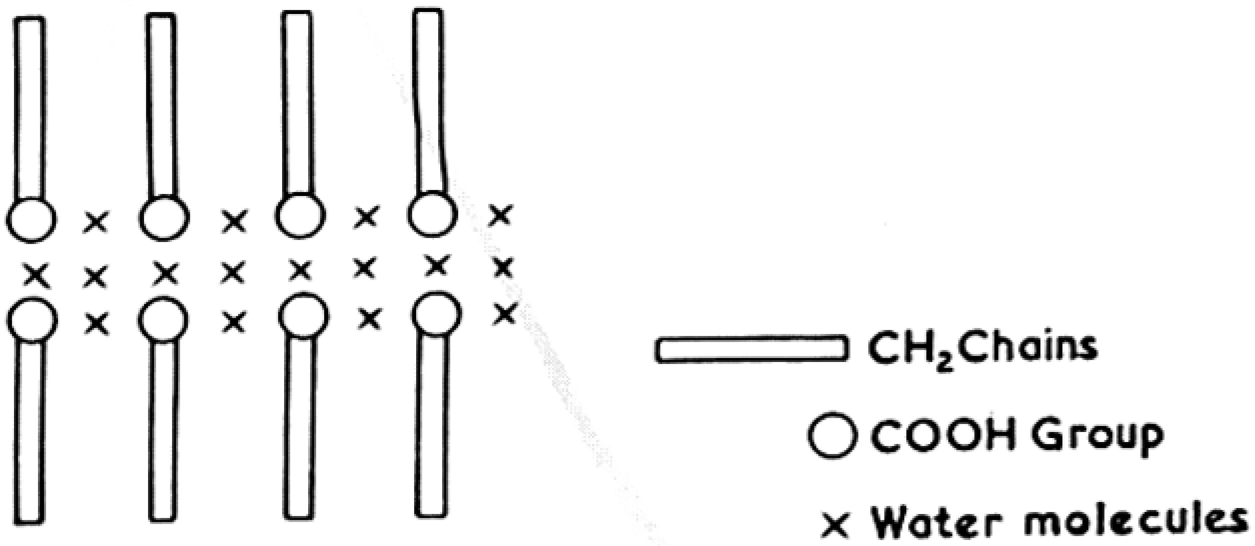


FIG. 92.

cases, such as sodium oleate, a film one molecule thick can apparently also exist, and, if so, there would be another region of negative dT/dt between S and O in Fig. 91.

(*b*) *Surface tension of pure liquids.*—In much of the early work on surface tension, the importance of really clean conditions was not realised. It is now known that a film of grease as little as one molecule thick can have a noticeable influence on surface tension. Thus, if the oil molecules are of the order of 10^{-6} cm. long, 1 c.c. of oil would cover 100 square metres of water with such a film. It is the fact that such a thin film produces a noticeable effect on surface tension that enables waves at sea to be stilled by quite moderate quantities of oil. The damping effect is believed to be due, not so much to the wind being prevented from "taking hold" of the water, but to the fact that, if the surface tension is a function of the thickness of the oil layer, the work done in stretching out a surface against surface tension will not necessarily all be recovered when the surface contracts again, as it would be with a pure liquid with fixed surface tension.

Plainly the weight of oil required to contaminate the surfaces of water in an ordinary piece of laboratory apparatus is very small indeed (of the order of a microgram), and could easily come from, *e.g.*, a finger touching the surface. It is a comparatively simple matter to clean grease from metal and glass apparatus; what is less easy is to remove it from the surface of the liquid. Mercury can be cleaned by chemical treatment, but almost the only way to remove grease from water is to "scrape" the surface with a piece of metal, cleaned or covered with paraffin wax, which piles the grease up in front of it, and leaves a clean surface behind.

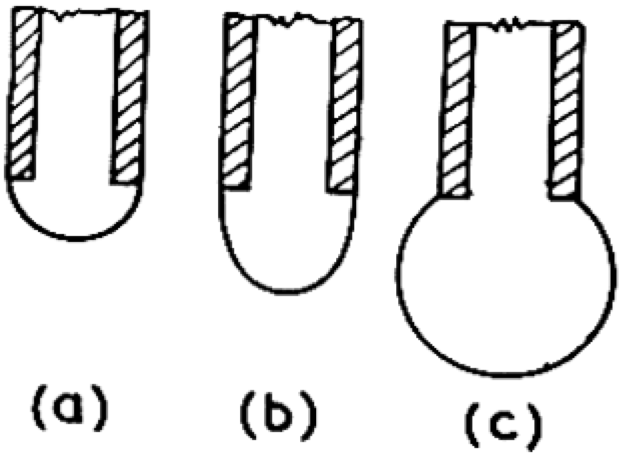


FIG. 93.

We have outlined above one of the simplest methods of measuring surface tension, the capillary rise method (p. 243) in which it is only necessary to measure the rise and the bore of the tube. As we have seen, this needs a measurement of the angle of contact by some independent method.

A second very direct method is due to **Jaeger**, who measured the excess pressure needed to blow bubbles in the liquid from a vertical jet. It is sometimes assumed that the bubble will break off as soon as the excess pressure corresponds to the value $2T/R$, where R is the outer radius of the point of the jet. For (Fig. 93) it is evident from geometrical considerations that in both Figs. 93(*a*) and 93(*c*) the radius of the bubble is larger, and the excess pressure inside it therefore smaller, than is the case for the perfect hemisphere, so that if we apply a constant pressure sufficient to produce a hemisphere it will expand to larger and larger radii until it becomes unstable and breaks off. In practice there is some uncertainty about the absolute value of T, because gravity distorts the bubble from the truly hemispherical shape. (Bashforth and Adams' tables can be applied here to compute the true shape of the bubble.) This method is therefore used chiefly as a rapid means of *comparing* surface tensions (particularly in investigating the change of surface tension with temperature). It has the advantage that a freshly formed liquid surface is usually cleaner than an old one.

Another very direct method, due to **Wilhelmy**, is to measure, by a balance or torsion balance, the force necessary to drag a vertical metal plate [Fig. 94(a)] away from the surface. When the plate is about to leave the surface the force acting on it, provided that the angle of contact is zero, is Tl, where l is the *total* length of the liquid line on the plate, and allowance is made, if necessary, for the liquid displaced by that portion of the plate below the original free surface. Some workers have used a horizontal plate, for which the theory is a little more complicated [Fig. 94(b)]. We consider the equilibrium of the portion of liquid whose section is $ABCD$. The horizontal forces acting on it are $2T$ per unit length acting along the lines through C and D (again provided that the angle of contact is zero), the atmospheric pressure acting over the arc CD, the corresponding horizontal force being $p_0 d$ per unit length, where p_0 is the atmospheric pressure and d is the depth AB, and the force acting across AB. The pressure at a depth y above B is $p_0 - g\rho y$, so that the total horizontal force across AB is

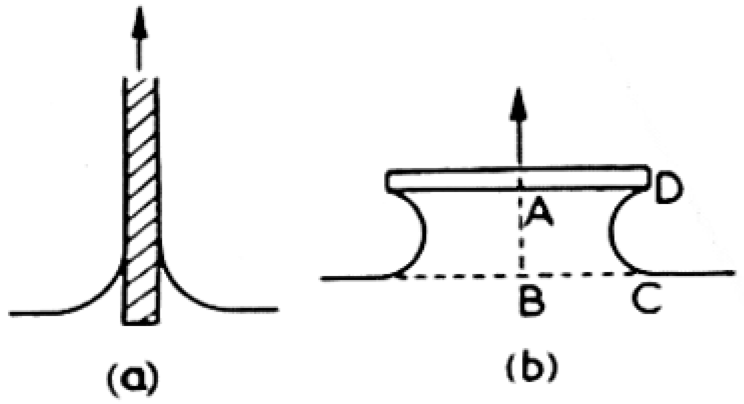


FIG. 94.

$$\int_0^d (p_0 - g\rho y)\,dy = p_0 d - \tfrac{1}{2}\rho g d^2$$

so that our condition for equilibrium is

$$\tfrac{1}{2}\rho g d^2 = T(1 + \cos\theta),$$

where θ is the angle of contact at D. The pressure-difference between the top and bottom of the plate is

$$\begin{aligned} p_0 - (p_0 - \rho g d) &= \rho g d \\ &= \sqrt{2T(1 + \cos\theta)\,\rho g} \quad \ldots\ldots\ldots\ldots \quad (21) \end{aligned}$$

so that the corresponding force is given by this pressure-difference multiplied by the area of the plate. Other workers have used a flat ring instead of a plate, for which the theory is very similar.

It will be noticed that we obtained the result without having to calculate the shape of the meniscus CD explicitly.

A method on a similar principle, which is due to **Quincke**, is to measure the dimensions of bubbles and drops. Let Fig. 95(a) represent a section by a plane through the axis of a drop of mercury on a glass surface. If the drop is large enough, the pressure in the liquid at A is the same as that of the atmosphere, if this is not so we can apply a correction $2T/R$, where R is the radius of curvature at A. The horizontal forces acting on a portion of liquid whose section is ABE and whose thickness perpendicular to the paper is dt consist of the hydrostatic pressure across AE, the corresponding force being $(p_0 l + 2Tl/r + \frac{1}{2}\rho g l^2)\,dt$, where $l = AE$, the force Tdt due to the surface tension along the line through A perpendicular to the paper, and the atmospheric pressure acting over the curved surface AB, the corresponding horizontal force being $p_0 l dt$, so that we have, from the equilibrium of the section ABE

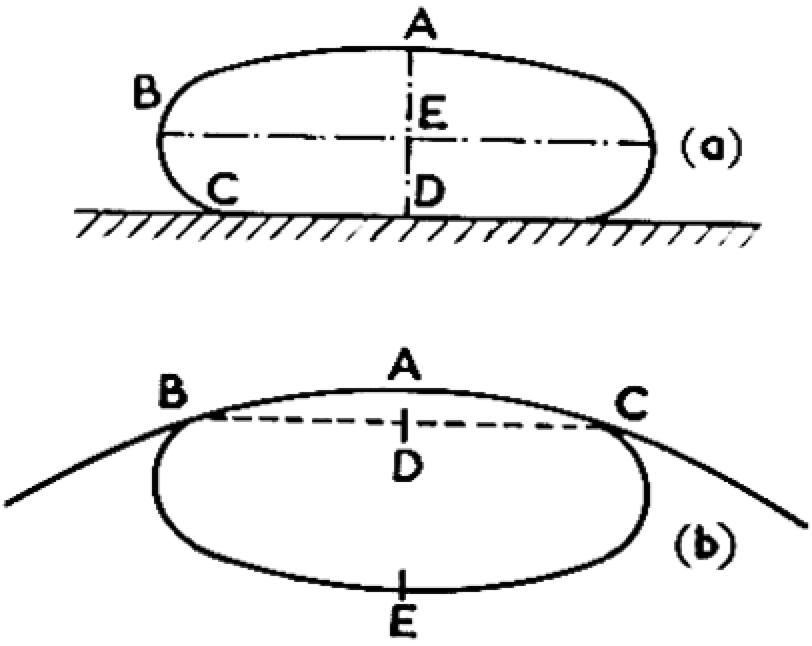


FIG. 95.

$$T = \frac{2Tl}{R} + \tfrac{1}{2}\rho g l^2 \quad \dots\dots\dots\dots\dots\dots \quad (22)$$

By considering the equilibrium of the portion of thickness dt and section $ABCD$ [Fig. 95(a)] we get, in a similar way,

$$T\,(1 + \cos\theta) = \frac{2Th}{R} + \tfrac{1}{2}\rho g h^2 \quad \dots\dots\dots\dots \quad (23)$$

where h is the distance AD, and θ is the angle of contact at C. In principle, these two equations enable us to determine both T and θ. In practice, the distance AE or l is hard to measure accurately, whereas h can be measured with great accuracy by a spherometer, and R is easily measurable, *e.g.* by an optical method.

For liquids which wet glass we can vary the method by using a bubble under a clock-glass [Fig. 95(*b*)]. We can apply formulae such as (22) or (23), but we now consider the equilibrium of sections such as $BDCE$, which means that h and l have to be measured from D and not from A. Further, θ has to be replaced by the angle the surface of the liquid at B or C makes with the horizontal. This is not quite the same as the true angle of contact. In practice, the points A and E can be located readily with a vertical travelling microscope, AD is determined geometrically from the curvature of the clock glass and the diameter BC of the circle of contact of the liquid with the glass. Here again, we are able to deduce the surface tension without a complete theory of the profile of the surface, which would be difficult to work out.

Another method, also used by **Quincke** and **Lord Rayleigh**, involves measuring the weight of a drop that forms slowly on the end of a tube (Fig. 79). One might at first sight think that we can write simply $mg = 2\pi aT$, where a is the outer radius of the tube, by considering that the drop is held up by the surface tension acting all round the circle of radius a. This, however, is only a rough approximation, and is in error for the following reasons:—

(i) Not all of the liquid shown in Fig. 79 forms the actual drop, some of it being left behind.

(ii) The boundary surface of the drop is not vertical at its junction with the tube.

(iii) In practice, the weight of the drop is determined by counting the number which fall in a given time, and then weighing the liquid. Thus, essentially, the experiment is a dynamical one, the complete theory of which is likely to be difficult.

The form of result to be expected if the rate of formation is kept slow can, however, be predicted by dimensional analysis (see Chapter X). Let m be the mass of a drop. We seek a relation between m and these quantities: the surface tension T, the acceleration of gravity g, the density ρ, and the outer radius of the tube a. The inner radius is unlikely to have any very great effect. We seek relations of the type $T^{\alpha}\rho^{\beta}g^{\gamma}a^{\delta}m^{\epsilon} =$ constant. The dimensions of the quantities are

$$T = \text{Force per unit length} = [M]\,[T]^{-2}.$$

($[T]$, as usual, represents the manner in which the unit of time enters, and is not to be confused with T.)

$$\begin{aligned} g &= [L]\,[T]^{-2} \\ m &= [M] \\ a &= [L] \\ \rho &= [M]\,[L]^{-3} \end{aligned}$$

By equating indices of $[M]$ $[L]$ and $[T]$ to zero, we conclude that the following two dimensionless quantities can be built up out of the above five variables:

$$Ta/mg \quad \text{and} \quad \rho a^3/m$$

so that the most general possible relationship involving only these variables, see p. 278, is

$$\frac{Ta}{mg} = f\left(\frac{\rho a^3}{m}\right) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (24)$$

The form of the function f was determined by **Harkins** and **Brown** by experiments on a number of liquids of known surface tension, and they further established that the form of the function f was the same for four very dissimilar liquids, water, benzene, carbon tetrachloride, and ethylene dibromide. Thus, the essential soundness of the assumptions on which the dimensional analysis was based seem to be confirmed, and the method can be used with confidence for measuring unknown surface tensions.

It is also possible to arrive, by dimensional arguments, at a relationship between the surface tension of a drop and its time of oscillation about the spherical form. A related, practically equivalent problem, is the relation between the wavelength of the varicose oscillations of a jet, and its radius, and the density and rate of flow of the liquid. For a detailed theory of the oscillations of a spherical drop under surface tension, the reader is referred to Lamb's *Hydrodynamics*, but the dimensional argument is here extremely simple. The period t, of oscillation of the drop, supposing that it is small enough for the effect of gravity to be neglected, is likely to depend only on the surface tension T, the radius of the drop a, and its density ρ, and, by seeking a relation of the type

$$t = \text{Constant} \times T^{\alpha}\rho^{\beta}a^{\gamma},$$

we find that the only possible one is

$$t = \text{Constant} \times \sqrt{\frac{\rho a^3}{T}} \quad \ldots\ldots\ldots\ldots\ldots \quad (25)$$

the complete theory giving the value $\pi/\sqrt{2}$ for the constant.

In practice, the time of oscillation of a drop, of radius small enough to exist as a sphere, *i.e.* for surface tension forces to be large compared with gravity, cannot easily be measured directly, but it can be determined indirectly by measuring the profile of a jet just before it breaks up (Fig. 86). Evidently the period of oscillation of a drop, of the size of the large drops into which the jet breaks up, is given by the ratio of the distance between maxima of the sinuosities of the jet and the velocity with which they advance and both these quantities can be determined, *e.g.* from a series of cinema pictures of the jet.

Another dynamic method is by the measurement of the speed of propagation of ripples. The theory of this is given on p. 166, the formula being

$$V^2 = \frac{2\pi T}{\lambda\rho} + \frac{g\lambda}{2\pi} \quad \text{.....................} \quad (26)$$

and for short wavelengths the second term is small compared with the first. In the method devised by **Lord Rayleigh** the ripples are produced by an electrically driven tuning fork of known frequency with a stylus dipping into the liquid. A shutter is driven in synchronism with the fork, so that the ripples are illuminated stroboscopically, and therefore present an apparently stationary pattern. The wavelength of the ripples can be measured by projecting their stroboscopic image on to a screen with a known magnification, and measuring the distance between maxima. The wavelength and frequency being known, V can be determined and substituted into equation (26).

(*c*) *The surface tension of contaminated liquid surfaces.*—We are very often interested in the effect of contamination of a liquid surface, *e.g.* by grease. Most of the static methods we have described above can be applied safely, but it would not be really sound to apply the dynamic methods we have described, or, at least, we should not necessarily expect them to give the same results as the static methods, because of the possibility that, *e.g.* in the ripple method, the grease film is of different thickness at the crests and troughs. However, in some of the static methods we have described, it might be a difficult matter to decide exactly how much contamination is actually present.

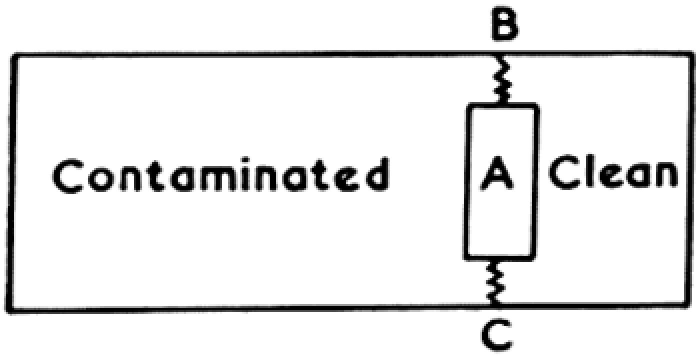


FIG. 96.

Accordingly, much of the modern work on surface contamination is based on a method due essentially to **Miss Pockels** and **Lord Rayleigh.** In principle, what is done is to work with liquid in a trough or framework whose area is exactly known, but with one side movable (Fig. 96). If we start with a clean surface, we can contaminate it with a known small quantity of grease dissolved in a solvent such as carbon-tetrachloride or benzene which evaporates rapidly. The area covered by the grease can be varied at will by moving the barrier A (Fig. 96) along the surface when it will heap up the grease on one side and leave a clean surface on the other. (Light ligaments of metal foil, B and C, can be used to prevent the escape of grease round the corners.) The force needed to push A to the left can be measured by, *e.g.*, a torsion balance. The force on A will almost always be to the right, which implies that the surface tension is diminished by contamination. We can also measure the surface tension within the barrier, *e.g.* by Wilhelmy's method. Fig. 96, however, suggests an alternative interpretation (equally valid) of the force on the movable side. Instead of saying that the surface tension of water is modified by contamination, we may equally well say that the observed force on *A* represents a two-dimensional analogue of pressure exerted by the film, due to the impact of grease molecules on the barriers. As we diminish the contaminated area, a number of interesting changes occur in the structure of the film. The force on the movable side is very small until there is enough grease to form a unimolecular film. Before this, the molecules are probably moving about the surface independently or in small clusters, and we may call such a film "gaseous". When a unimolecular layer is formed the surface tension decreases slowly. As we compress the film further, more and more layers are formed, until we finally reach a stage when the film behaves like a solid, and shows folds and buckles up instead of increasing in thickness further. The intermediate stage is often known as a "liquid" film by analogy with the properties of liquid in bulk, though some films show properties that have no exact analogues in three dimensions. The reader is referred to N. K. Adam's *Physics and Chemistry of Surfaces.*

(*d*) *The surface tension of liquid interfaces.*—A knowledge of the surface tension of liquid-liquid interfaces is essential when one is discussing, *e.g.*, the spreading of liquids over one another. Most of the methods that we have described for the measurement of the surface tension of a pure liquid are, in principle, applicable here also, for example, one can use the weight of drop method, the drops being formed by one of the liquids in the other, instead of in air. Again, we can measure the dimensions of a drop of one liquid formed underneath the other. It is necessary to modify the formulae appropriately to allow for the finite density of the liquid within which the drop is formed, but this usually presents no difficulty, and simply requires that ρ, the density, should be replaced by the difference of the two densities. If we prepare two liquids of nearly equal densities we can study such phenomena as the formation of drops and the break-off of the "neck" on a very large scale. The densities of aniline and water are very nearly the same, aniline being slightly denser at room temperature, while it is slightly lighter than water near 100° C. Thus, if a hot mixture of aniline and water is allowed to cool, the phenomenon of the formation of drops can be closely studied, the drops being very large and forming fairly slowly, if the temperature is maintained just below that at which the densities are equal, because the order of magnitude of the surface tension is the same as for an air surface, but the effect of gravity is much diminished.

(*e*) *Direct measurement of angles of contact.*—It is possible to estimate an angle of contact approximately from a comparison of two methods of measuring the surface tension, one of them independent of the angle of contact, such as the weight of drop method, the other dependent on it. However, the results are always rather uncertain, and it does seem that there is a significant difference in the measured angle, according to whether it is measured in advancing or receding conditions, and at what relative velocity.

The first direct measurement seems to have been due to **Gay-Lussac,** who gradually filled up a spherical glass flask with mercury until the meniscus was accurately plane. This can be judged by eye very accurately. A very simple method was used by **Adam,** who placed the material in the liquid in the form of a plane slab, the angle being continually adjusted until the neighbouring surface was flat. A more refined method was to lower a cylinder progressively into the liquid until the meniscus, viewed by an optical method, was seen to be accurately flat. By rotating the cylinder slowly, it was possible to detect any difference between the advancing and receding contact angles, and to investigate the effect of changing the relative velocity. The results revealed a fairly smooth dependence on the velocity, but showed clearly that angle of contact is something that depends on many factors, and, like such a quantity as the magnetic permeability of iron, cannot always be made to assume a precise value merely by refining the methods and precautions of measurement.

5. The spreading of liquids over one another. Surface tension of solutions

In Fig. 97 we see the situation to be expected if we have one liquid that does not spread over the other. The line through O, where the three liquids meet, is subject to the three forces, T_1, T_2, T_{12} per unit length, corresponding to the surface tensions of the free surfaces of liquids 1 and 2 and the surface tension of the interface, respectively. If such equilibrium were possible, we should be able to draw a triangle of forces (Fig. 97), the three sides proportional to the three tensions. Such a triangle is called *Neumann's triangle*, but few, if any, cases seem to be known in which it can be drawn, in other words, one of these tensions is nearly always larger than the sum of the other two. This means that one liquid nearly always spreads over the surface of the other.

Some of the higher paraffins, such as are found in crude petroleum, do seem to remain on the surface of water in the form of lenses, without spreading over the surface, but with other liquids, such as benzene, which *appear* to behave in this way, what has really happened is that a very thin layer of benzene has first spread over the water, and this thin layer can then remain in equilibrium with the rest of the benzene, which collects on the surface in the form of lenses. **Hardy** distinguishes between *primary spreading*, the initial formation of the thin layer which may occur by flow along the surface, or via the vapour phase, and *secondary spreading*, the spreading of one liquid to form a uniform layer over the other. Other cases in which there is apparently no spreading, such as water placed on the surface of mercury, are accounted for by the contamination of one of the surfaces. If mercury is really free of grease, water spreads quite easily over it.

Surface tension of solutions.—An empirical rule which has been verified in quite a number of cases is due to **Antonow**. It states that the condition $S_{12} = |S_1 - S_2|$ holds for two liquid phases which are not pure but are saturated solutions of each liquid in the other (and are therefore in equilibrium with one another, and with the same vapour phase). This empirical rule may have some theoretical significance, but this has not yet become clear. The surface tension of solutions of two liquids miscible in all proportions usually follows a simple course between the surface tensions of the two components. Sometimes there are well-defined maxima or minima, or points of inflection, but these are usually at concentrations corresponding to the formation of definite chemical compounds, *e.g.* sulphuric acid is known to form various hydrates as well as dissolving in water. The solution of inorganic electrolytes in water generally raises its surface tension and the same is true of some organic substances such as sugars and glycerine. On the other hand, alcohols, aldehydes, and fatty acids lower the surface tension. This is particularly noticeable with camphor, and results in the well-known phenomenon of the rapid motion of small pieces of camphor on a clean water surface.

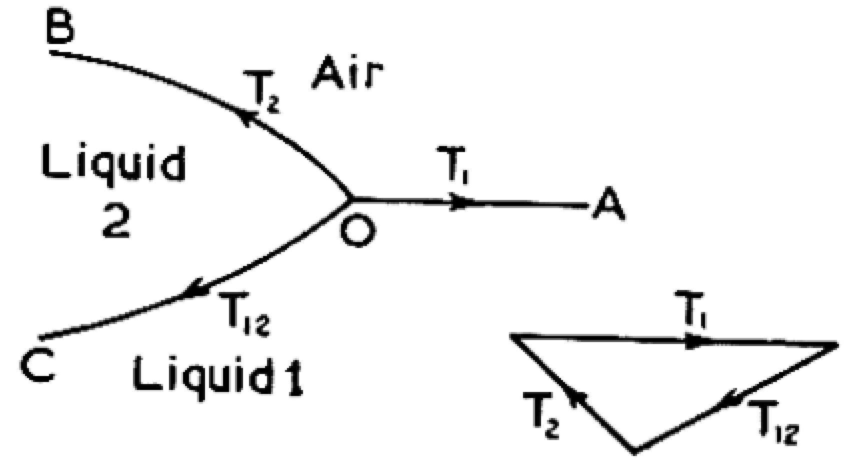


FIG. 97.

The results of measurement of surface tension of solutions are hard to interpret theoretically, because, not only is it almost certain that the density of a pure liquid varies rapidly in the neighbourhood of a surface, but we also have to consider the possibility that the *concentration* of solute may also change here. Indeed, if we have a solute that increases the surface tension, the surface energy will be decreased if we transfer some of the solute from the surface layer to the body of the liquid. Conversely, a solute that *lowers* the surface tension would tend to reach a higher concentration in the surface layer. (This of course does not mean that nearly all the solute will collect in the surface layer, because such an occurrence would be prevented by diffusion.) The equilibrium between the surface layer and a bulk solution was considered thermodynamically by **Gibbs,** and the reader is referred to textbooks of chemical thermodynamics for details.

The above reasoning shows, and experiment confirms, that dissolving a substance in a liquid may *lower* the surface tension very appreciably but cannot *raise* it to any great extent because raising the surface tension weakens the concentration near the surface. So-called "wetting agents" are usually molecules chemically similar to soaps, and they really act by causing a marked reduction in the surface tension of water to values of the order of 25 dynes/cm., as compared with the value of over 70 dynes/cm. characteristic of pure water, thus making it easier for the liquid to enter narrow channels and cracks.

6. The Laplace theory of surface tension and allied effects

We shall now develop the consequences of the two main assumptions of **Thomas Young** and **Laplace,** namely, that a liquid-vapour boundary can be safely treated as abrupt and that the forces between molecules may be considered to be of short effective range. The assumption of an abrupt boundary is certainly not correct, and **van der Waals** made attempts to improve it. It seems likely that the transition is virtually complete within a few atomic distances, so that the inaccuracies introduced by the cruder assumption are probably not serious.

The second postulate is qualitatively correct, for it is known that, for molecules without a permanent electric moment, the forces between them fall off at least as fast as the inverse seventh power (p. 188). For the present, we shall leave the precise form of the interaction unspecified. We suppose that the attraction of an infinite flat plate of liquid on a molecule at a distance z above it is given by the function $\psi(z)$ (Fig. 98). This function will be practically independent of the thickness of the plate, as long as it is large compared with the effective distance of molecular interaction, which we may denote by c. We now calculate the force per unit area acting across the plane dividing the two portions

A and B of the liquid. The force exerted by A on a layer of thickness dz is $N\psi(z)\,dz$ per unit area, where N is the number of molecules per unit volume. Thus, the force exerted by A on the whole portion B is

$$N\int_0^\infty \psi(z)\,dz \text{ per unit area.}$$

This we may identify with the correcting term a/V^2 in the van der Waals' equation, the so-called "intrinsic pressure". We now ask what is the work required to pull away a plate of liquid of small thickness δ from the liquid, and carry it away to infinity. This is

$$N\delta\int_0^\infty \psi(z)\,dz \text{ per unit area.}$$

Now, for a thin plate, the work necessary to remove one molecule from it is proportional to the thickness of the plate, that is to δ, so that the work done in breaking up this plate into individual molecules is proportional to the number of molecules, and to the work done in removing one, that is to δ^2, which may be neglected in comparison with δ. We thus get a second interpretation of the integral of $\psi(z)$, namely, this integral, in suitable units, represents the latent heat of evaporation.

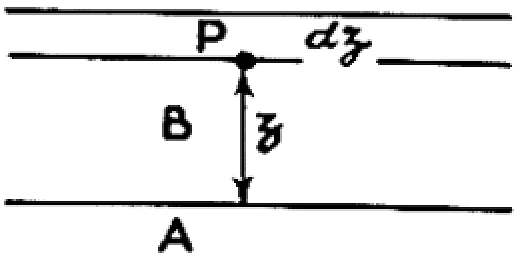


FIG. 98.

A third interpretation of the function $\psi(z)$ can be obtained as follows. We suppose the whole of portion B of the liquid removed to infinity, against the attraction exerted on it by A. The result of such an operation is simply to create two new liquid surfaces. We divide up the liquid B into horizontal slices as before, and suppose them removed one by one to infinity against the attraction of A. We need not consider the attractions *between* the various slices of B, because the work done against these is recovered when the various slices are joined together again at infinity. Thus the work done in removing the slice dz to infinity against the attraction of A is

$$Ndz\int_z^\infty \psi(x)\,dx \text{ per unit area.}$$

So that the work done in removing the whole of B out of the influence of A is

$$N\int_0^\infty dz\int_z^\infty \psi(x)\,dx \text{ per unit area.}$$

Integrating by parts, we see that

$$\text{Work/Unit area} = N\Big|_{z=0}^{z=\infty}\left[z\int_z^\infty \psi(x)\,dx + \int z\psi(z)\,dz\right].$$

Now, the first term vanishes at $z = 0$, and also at $z = \infty$ (because in the latter case the upper and lower limits of the integral are the same), so that we have finally

$$\text{Work/Unit area} = N\int_0^\infty z\psi(z)\,dz$$

and the surface tension is just half this because two new surfaces of unit area are created by this work, so that our final results are

$$T = \frac{N}{2}\int_0^\infty z\psi(z)\,dz, \quad \frac{L}{V} = N\int_0^\infty \psi(z)\,dz \quad \ldots\ldots\ldots \quad (27)$$

These results were the basis of the historic attempt by **Thomas Young** to estimate the range of intermolecular forces, for, if we suppose that $\psi(z)$ is inappreciable for z greater than the value c, but positive for all smaller values, we arrive at the conclusion

$$T \leqslant \frac{N}{2}c\int_0^\infty \psi(z)\,dz$$

or

$$c \geqslant \frac{2TV}{L}.$$

For water we have $T = 75$ dynes/cm. and $L = 600 \times 4{\cdot}2 \times 10^7$ ergs/g., so that $c \geqslant 6 \times 10^{-9}$ cm., while we should get very similar values for other liquids. This estimate is of the correct order of magnitude. Another important result that Laplace deduced was that the assumption of a definite "cut-off" of the function $\psi(z)$ at $z = c$, enabled him to prove that the excess pressure inside a curved surface of radius R could be represented by the expression $2T/R$ that we obtained by other means on p. 237.

We now believe that intermolecular forces fall off according to a power law, probably r^{-7} at large distances (p. 188), which, as we shall see in a moment, leads to the conclusion that $\psi(z)$ should behave as z^{-4}, which cannot be represented properly by any function with a finite cut-off. We therefore omit Laplace's proof.

To make further progress, we must assume some definite form for $\psi(z)$. To do this we suppose that the slab of liquid A contains N molecules per unit volume, arranged in a completely random manner, and we calculate $W(z)$, the mutual potential energy of this slab and a single molecule at a distance z. We suppose the mutual potential energy of two molecules to be $-C/r^6$ for $r > r_0$ and positive infinite for $r < r_0$. As we have seen (p. 188) there is theoretical justification for the inverse sixth power law of attraction, while it is clearly necessary to introduce a repulsive term to explain the stability of liquids and solids. Thus, our model of "attracting rigid spheres" is the crudest that has any hope of representing physical reality.

To calculate $W(z)$, we divide the slab A up (Fig. 99) by cylinders whose axes pass through the molecule at a distance z from A. Consider now an element of volume cut off between two such cylinders of radii x and $x + dx$ and lying between two planes at distances y and $y + dy$ below the top surface of A. The volume of such an element is $2\pi x\,dy\,dx$, and the distance r of each point in it from P is nearly equal to $[(z + y)^2 + x^2]^{\frac{1}{2}}$, so that the mutual potential energy of the molecules within the element of volume and the molecule at P is

$$-\frac{2\pi C N x\,dy\,dx}{[(z + y)^2 + x^2]^3}\quad z > r_0,$$

since there are N molecules per unit volume in the slab. Integrating, we have for $W(z)$ the result

$$W(z) = -2\pi CN\int_0^\infty\int_0^\infty \frac{x\,dx\,dy}{[(z+y)^2 + x^2]^3} = -\frac{CN\pi}{6z^3}\quad z > r_0$$
$$= +\infty\quad z < r_0.$$

Differentiating, we have for $\psi(z)$ the result $CN\pi/2z^4$ $(z > r_0)$, it being understood that the molecule cannot approach closer to the slab than r_0, so that integrals involving $\psi(z)$ are to be taken with this as a lower limit. Substituting this result in equation (27) we find

$$\frac{L}{V} = \frac{CN^2\pi}{6r_0^3}\qquad T = \frac{CN^2\pi}{8r_0^2}.$$

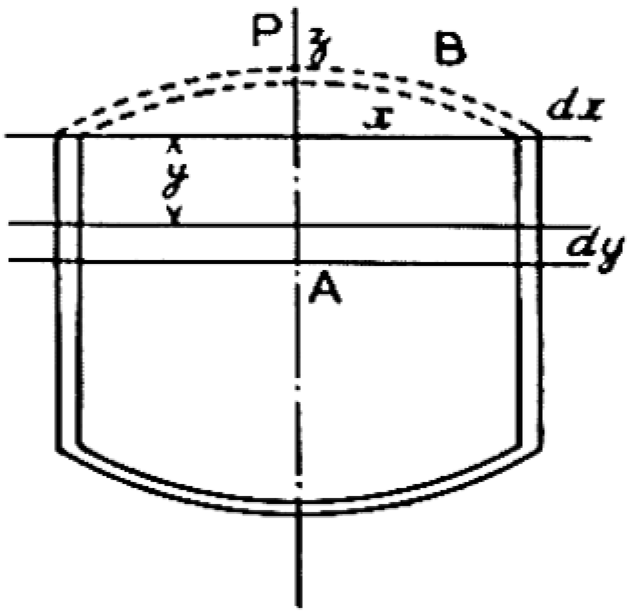


FIG. 99.

This time, by comparing L and T, we get an estimate, not for c, the range of the *attractive* forces, as **Young** supposed, but for r_0, the radius of the "hard core" of our molecules. For most liquids TV/L is of the order of 10^{-8} cm., which is in accord with what is known about atomic diameters. **Hammick**, by a more elementary argument, deduced a similar proportionality between L/V and T/r_0, but with a different numerical factor, and showed that it is obeyed by a number of substances. However, any such relation is a consequence of the very simple law of force we have assumed and would not hold in the more general case, *e.g.* if we supposed that the repulsive field is also governed by a power-law.

It will be noticed that our final result for L/V contains the factor N^2, so that the Laplace theory of surface tension really leads to an approximation of the same type as that obtained by van der Waals for a gas (p. 191), in which the correction to the pressure is also proportional to N^2. Our work clearly shows the origin of such a variation—when calculating $W(z)$ or $\psi(z)$ we supposed the molecules in the slab to be distributed in random fashion. Actually, the approach of our single molecule must modify this distribution somewhat, attracting the molecules in the slab towards itself, and this effect would give rise to correcting terms involving higher powers of N, as are known to be present in a dense assembly.

Another serious criticism is that the changes in temperature resulting from the processes considered by Laplace ought to be taken into account if we imagine them carried out adiabatically. It is therefore not clear whether $\psi(z)$ is related to surface tension, surface energy, or to some intermediate function.

7. Variation of surface tension with temperature. The parachor

The surface tension of liquids invariably decreases as the temperature rises. This is because, as we approach the critical temperatures, the liquid and vapour phases become more and more alike. The decrease is usually fairly smooth and slow, so that it can be represented by a linear law over considerable ranges of temperature. A number of empirical formulae have been proposed to represent this variation, but the only one that we shall consider is that due to **Macleod.** This is

$$T = C(\rho_l - \rho_v)^4$$

where C is a constant for each substance and ρ_l and ρ_v are the densities of liquid and vapour respectively. This relation was shown by **Fowler,** by a further development of the Laplace type of theory of surface tension, to be theoretically of the form to be expected in the immediate neighbourhood of the critical point. However, he was unable to say over what range of temperatures such a relation might be expected to be valid, so that the fact that it is experimentally valid at room temperature must still be regarded as empirical.

This relation is of interest from another point of view, since it led **Sugden** to argue that, since M/ρ_l might be regarded as a rough measure of the volume of a molecule of substance, M being the molecular weight, the quantity $MT^{\frac{1}{4}}/(\rho_l - \rho_v)$ might be regarded as a better measure of molecular volume since it is independent of temperature over a fairly wide range. This volume is to be computed for each substance at a temperature such that T takes a standard value (say unity). This quantity, expressed in suitable units, was named by Sugden the *parachor*, and he was able to show that, if suitable values were assigned to each kind of atom, the parachor for a molecule could be predicted with good accuracy merely by adding up the values for the constituent atoms. In certain cases, such as the existence of, *e.g.*, double bonds of ethylene or of semi-polar type, a triple bond, or a benzene ring, it was necessary to make a special allowance over and above the contributions of the individual atoms, but it was shown that this could nearly always be done consistently. Thus the parachor could be used as a means of deciding whether or not certain special types of structure are present in an organic compound and thus of deciding between various proposed structural formulae.

The physical meaning of this speculation is by no means clear, but its successes were too striking to be dismissed as purely accidental. The method would hardly be used to-day as a means of deciding between two or more possible structures of an organic compound, because at least two better and more precise methods are now available, namely X-ray analysis, which has now been brought to a stage at which characteristic features such as the benzene ring can actually be mapped out and recognised, and infra-red spectroscopy, for it has been shown that certain absorption bands are associated with definite structural features of an organic molecule.

8. The concept of association in a liquid

We developed the Laplace theory of surface tension on the basis that the molecules of a liquid could be regarded as independent units, with interaction forces of "central" type, that is to say, acting between the molecules as a whole rather than between specific parts of the molecule. Now it has long been known that such an idea cannot be entirely correct, for example, it cannot be correct for molecules such as ketones or nitrobenzene, which possess localised dipoles by virtue of their $C{=}O$ or NO_2 groups, which would result in a localised directional type of interaction that cannot be represented by a force between the centres of gravity of the molecules. Later on, the discovery of "hydrogen bonds" showed that neighbouring molecules could actually be linked together by a hydrogen atom between, *e.g.*, two oxygen atoms.

Gradually a distinction has come to be drawn between "associated" and "non-associated" liquids, mainly on empirical grounds. The physical distinction between the two is quite sound, but it is a matter of historical fact that it was arrived at in an extremely tentative way. Liquids were considered to be "normal" if they obeyed certain empirical rules, for most of which no theoretical basis was known, nor is any known now. Among such rules we may mention *Trouton's rule* (that the latent heat of vaporisation of many liquids is nearly the same multiple of its boiling point), the empirical relation of **Richards** and **Matthews** (that $T\beta^{\frac{4}{3}} = \text{Constant}$, where β is the compressibility), and the observation of **Bennet** and **Mitchell** [that surface *energy*, as distinct from surface *tension* (see p. 232), is constant over a wide range of temperatures]. Among arguments of a similar "intuitive" type we may mention the fact that water and the alcohols have very high boiling points, compared with other liquids of equal molecular weights, which again suggests that such liquids are behaving as if they were composed of double or treble molecules.

We can only put it on record that the concept of "association" was arrived at in this empirical way, and remark that these intuitive arguments proved satisfactory. It is possible to divide associated liquids into types, according to the number of points on each molecule at which these effective "links" can be formed. If there is only one, as in the ordinary alcohols, the process of association must stop when we have double molecules that can perhaps be broken apart fairly easily. If there are two such points, then we should expect the molecules to form either rings or open chains. The ring-like aggregates will behave like large molecules, but the open chains may lead to high viscosity. (The increase in viscosity of liquid sulphur when heated is probably due to the break-down of rings into open chains.) The third possibility is that there may be *more than two* possible points on a molecule where links can be formed, in which case the linkages will bind the whole liquid loosely together, the classic example of such a liquid being water. In ice there may be as many as four hydrogen bonds per molecule, in liquid water the average number varies between two and three. Thus there is no strict meaning in speaking of the "effective molecular weight" of water, or in thinking of it as made up of "double" and "treble" molecules.

9. The surface tension of solids

The development of the Laplace theory of surface tension of liquids, and the fact that it leads to acceptable physical conclusions, prompts us to ask whether similar calculations cannot be made for solids. Such calculations have been attempted for the alkali halides, and for some other crystals which have a simple structure, so that the mathematics do not become too difficult. They lead to the conclusions that the surface tension of a solid is undoubtedly a real phenomenon; that numerical values of hundreds, or even thousands, of dynes per centimetre are quite possible, and that the effects might appear in various ways, such as a contraction of the lattice distance near the surface. Evidence for such an effect has, in fact, been found in the alkali halides and in magnesium oxide, in very thin metal foils, such as silver and gold, and in the cleavage of thin sheets of mica (the lattice distances of which sometimes change after cleavage), and it has even been possible to obtain a few estimates of surface tension by such means, but results are scanty.

It is as well to point out at this stage that the surface tension of solids has nothing directly to do with the existence of a finite angle of contact between liquid and solid and no analogy with Neumann's triangle is possible. For (Fig. 100) the angle of contact is evidently determined by the equilibrium of a typical molecule at P, under the combined action of the surface tension of the liquid T_l, and of the attractions of the clean part of the surface PR and of the part PQ, covered with liquid. The surface tension of the solid itself may possibly have altered the lattice structure near the surface, but it cannot exert any direct influence on the contact angle. It is, however, quite possible for the apparently dry part of the surface, PR, to be covered by an invisible film of liquid.

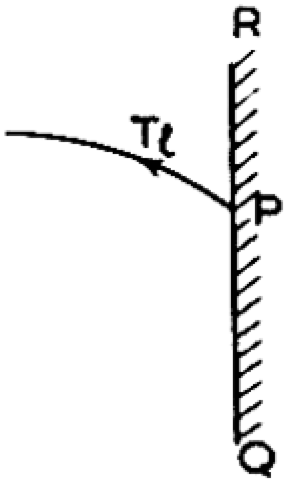


FIG. 100.

10. Adsorption of gases and liquids on solids

We have already pointed out that the properties of liquids can be profoundly altered by the presence of monomolecular layers of certain substances. It has become clear in recent years that the same is true of solids, and, in some cases, with even more startling results. It has been known for some time that the thermionic emission of electrons from heated metals is greatly affected by the presence of such films, but more recently, it has been found that the frictional properties of metals are also very much affected. Two pieces of metal under ordinary laboratory conditions slide over one another with a certain amount of friction, but do not "seize up" unless they get very hot. On the other hand, if clean metal surfaces are brought together in a vacuum, they usually adhere most tenaciously, and, if they are forcibly pulled apart, portions of each surface are dragged away. Freshly formed surfaces of lead will "seize up" even in air. Again, **Roberts**, investigating the behaviour of certain metals, found that their accommodation coefficient (p. 134) was very much smaller when the metal was clean than when it was covered with one or more layers of gas, and, indeed, that this fact could be used to follow the process of building up such layers.

Adsorption of gases by solids is the main factor in the performance of solid *catalysts* such as finely divided metals. In a few cases, the effect of catalysts can be accounted for by the formation and break-up of loose chemical compounds, and in one or two others, such as spongy platinum, hydrogen can not only be adsorbed on the surface, but it can form a "solution" in the interior as well, but in most cases, the reactions almost certainly take place on the solid surfaces themselves. One particularly interesting case is the adsorption of hydrogen. On most solids there are a definite number of "adsorption sites" on which an atom or molecule of the gas may rest. For hydrogen, on some metals, the attraction is so strong that the hydrogen molecule breaks up and is adsorbed as two separate atoms, being unable to adsorb as a molecule because its length does not correspond with the distance between two adsorption sites. This means that an adsorbed layer of hydrogen consists, in effect, of free atomic hydrogen, which is very much more active chemically than ordinary hydrogen. The "poisoning" of catalysts, caused by traces of impurities such as arsenic, is brought about by the adsorption of impurity atoms, which prevents the adsorption of atoms of the reacting gases. Very small quantities of impurity may be sufficient to cover all the active sites.

The theory of adsorption on a solid has been developed by a number of workers. For an adsorbed film to form at all, it is necessary that the binding energy between the solid and an adsorbed molecule or atom should be several times kT, where k is Boltzmann's constant, otherwise the film would be thrown off by thermal agitation. (Indeed, the most effective way of removing adsorbed layers from a metal is to heat it in a vacuum to as high a temperature as possible.) An energy of kT per molecule corresponds to approximately 600 calories per mole at room temperature. Adsorption energies on a clean surface are usually of the order of thousands of calories per mole, so that adsorption is very sensitive to temperature. A well-known example of this is the behaviour of charcoal, which adsorbs enormous quantities of gas at room temperature, while, if cooled by liquid air, it continues to adsorb gases down to pressures of a small fraction of a millimetre of mercury, so that it is used extensively in the production of high vacua. On the other hand, nearly all this adsorbed gas can be removed by heating the charcoal to redness. This naturally leads to the study of the relations between the quantity of gas adsorbed on a surface, and the pressure and temperature.

We may distinguish between two types of film, one being the so-called mobile film, in which the adsorbed atoms or molecules are bound to the surface, but are nevertheless free to move parallel to it. This may be thought of as a two-dimensional analogue of a liquid or gas. **Langmuir** suggested that, on any crystalline surface such as a metal, there are definite positions in the lattice at which gas atoms can be adsorbed easily, these adsorption sites lying between the atoms of the crystal. On any given surface, we may suppose that there are a definite number of adsorption sites, and that the film is complete when they are nearly all filled. (It may then be possible for a second layer to form on top of the first, but we shall not consider this question here.) Thus we distinguish between *mobile* and *immobile* films according to the behaviour of the adsorbed atoms actually on the sites. Once an atom has been adsorbed it may be an easy matter for it to travel to a neighbouring site, or, on the other hand, it may be so tightly bound to that site that the only processes that we need consider are the adsorption of an atom or molecule directly from the gas phase and the converse process of evaporation of the adsorbed atom from the site back into the gas phase. A film of this type we describe as *immobile*.

An adsorbed film may be thought of as an assembly of occupied and empty adsorption sites, and can be treated theoretically by an appropriate adaptation of a theory that has been applied to many other problems, such as binary solutions and alloys, to the "hole" model of liquids, and to problems such as magnetism, where we need only consider a comparatively small number of energy-levels associated with each atom. The simplest assumption that we can make in the adsorption problem is that the adsorption energy on a site is independent of whether neighbouring sites are occupied or empty. We can then derive the dependence of θ, the fraction of adsorption sites occupied, on the pressure and temperature of the adsorbed gas by a very simple argument, due to **Langmuir**. For, if we have a gas at a pressure P and temperature T, the number of molecules of mass m striking unit area of the solid per second is $P(2\pi mkT)^{\frac{1}{2}}$ (p. 125). The number that are adsorbed per second is proportional to the product of this and the number of vacant sites $N(1-\theta)$, where N is the total number of sites. The number that leave the filled sites per second is proportional to $N\theta$, and otherwise depends only on the temperature and the adsorption energy. We get equilibrium when the number of atoms that evaporate from the filled sites is equal to the number that are adsorbed on the empty sites per second, which gives us

$$X(T)\,\theta = Y(T)\,(1-\theta)\,P.(2\pi mkT)^{\frac{1}{2}}$$

where $X(T)\theta$ is the rate of evaporation of molecules from filled sites, while $Y(T)$ is the probability that a molecule striking an empty site will remain there. We may thus conclude that

$$\frac{\theta}{1-\theta} = PA(T) \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (28)$$

where $A(T)$ is a function of the temperature. **Fowler** showed that this equation, and the form of $A(T)$, can be deduced from statistical mechanics without having to resort to the "kinetic" type of argument. Equation (28) is usually known as *Langmuir's adsorption isotherm*, and it applies equally to mobile and immobile films.

To test this equation it is necessary to measure θ, the fraction of sites adsorbed, as a function of gas pressure. This can be done in various ways, of which we may mention the straightforward one of admitting known quantities of gas to the apparatus in successive small steps, and comparing the actual pressures attained with those to be expected if all the gas remained unadsorbed. Other methods, described in detail by **Roberts,** consist of measuring the accommodation coefficient (p. 134), of an inert gas such as neon, which is very much smaller on a bare surface than on one covered with adsorbed gas, and of measuring the Volta contact potential difference between the metal being studied and a filament being run at a high temperature, so that it may be considered to be bare. It is also possible to measure the adsorption energy for various values of θ. This can be done by having the adsorbing metal in the form of a wire and measuring the rises in temperature (by the change in its electrical resistance) as successive charges of gas are introduced. The reader is referred to Roberts, *Some Problems in Adsorption* for details.

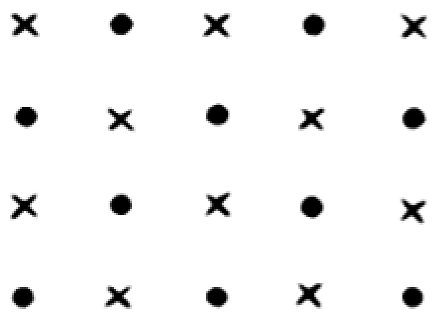


FIG. 101.

Analysis of such results shows that the Langmuir theory is often inadequate to explain the facts, because, as one might expect, there is usually a definite *interaction* energy between adsorbed atoms on neighbouring sites, which may be either attractive or repulsive, in addition to the energy gained through the attraction between the atoms and sites. If we take the repulsive case first (atoms disliking being on neighbouring sites) it will be seen that we can quite easily decide whether our film is mobile or immobile, by measuring the adsorption energy as a function of θ. For, if we have a mobile film and the adsorption sites are arranged in a square lattice (Fig. 101), the state of lowest energy for $\theta = \frac{1}{2}$ will be that in which occupied sites are represented by crosses. Thus, for $\theta < \frac{1}{2}$ it is always possible to find arrangements for which no two neighbouring sites are occupied. Thus, in this region, the adsorption energy should be nearly independent of the number of sites occupied. Once we have $\theta > \frac{1}{2}$, we must begin filling some of the sites represented by dots in Fig. 101, so that the effect of interactions between neighbours now has to be considered. Thus, we might expect an abrupt change in the measured adsorption energy as θ passes through the value $\frac{1}{2}$, this energy jumping from the value appropriate to a site whose neighbours are unoccupied to that for a site whose neighbours are all occupied. A more careful discussion shows that thermal agitation tends, as one might expect, to "round off the corners" of this predicted discontinuity.

For an *immobile* film, the predicted behaviour is entirely different. The molecules of gas strike the metal in random fashion, and we suppose that they either leave the metal or remain on the sites they strike. Thus, the arrangement of adsorbed atoms and empty sites is a random one for all values of θ, and this means that the probability that any specified site is occupied is proportional to θ, so that, if we consider the group of four sites surrounding any particular site, the expected number of neighbours is 4θ, and the interaction energy when a new atom is adsorbed is also proportional to θ. Thus, for an immobile film, the measured adsorption energy should decrease linearly as θ increases. This is what is found in practice for films of hydrogen on tungsten and enables one to conclude that such films must be regarded as immobile.

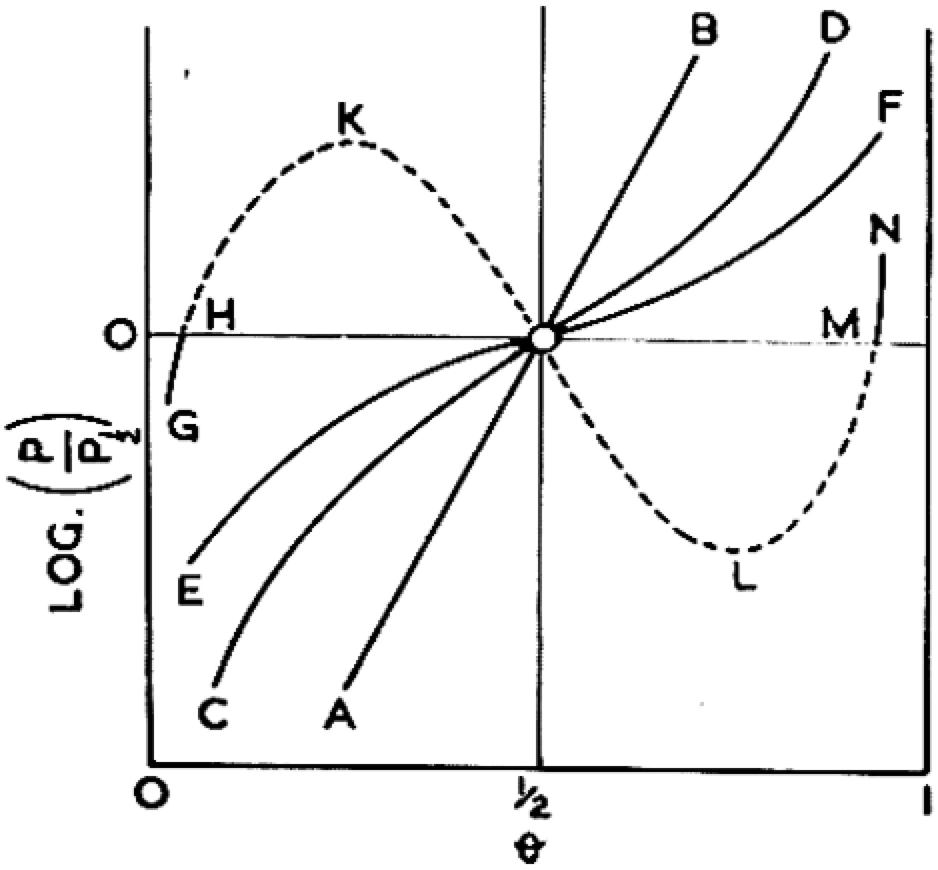


Fig. 102.

Since the interaction energy between a gas atom, or molecule, and the solid surface must be fairly large for the film to be stable at all, one would naturally expect the immobile type of film to be the most common. This seems to be so in practice, though certain large organic molecules form mobile films, and so does oxygen on tungsten at high temperatures.

The alternative assumption of *attractive* forces between adsorbed atoms leads to results of equal interest. In Fig. 102 we sketch **Fowler's** theoretical curves of $\log p$ as a function of θ, for various values of the ratio w/kT, where w is a measure of the change in potential energy caused by bringing two adsorbed atoms from non-neighbouring to neighbouring sites. (w positive implies repulsive forces, w negative attractive forces. w zero corresponds to Langmuir's adsorption isotherm.) In all cases we let $P_{\frac{1}{2}}$ be the pressure corresponding to $\theta = \frac{1}{2}$. Then curve AOB represents the case of w positive, COD Langmuir's adsorption isotherm, and EOF and $GHKOLMN$ two curves for which w is negative. Here we have the possibility of three values of θ corresponding to the same value of P, which will occur for sufficiently low temperatures, in just the same way as in the van der Waals' gas. EOF, which has a horizontal tangent, corresponds to the van der Waals' critical isotherm. For lower temperatures than this we deduce the result that if the pressure is slowly increased from zero, θ will at first assume small values and follow the variation of P, and will then, at a certain value of P, suddenly jump to a value of nearly unity. Further investigation shows that the equilibrium isotherm is $GHOMN$, but that the portions HK and ML are metastable and can be actually realised if the experiments are carefully done.

This phenomenon is called *critical adsorption* and was observed by **Cockcroft** for cadmium vapour adsorbed on copper. As the vapour pressure is increased, the adsorption is at first small, but, at a certain critical pressure, the first layer is nearly filled up and this permits further layers to form so that the cadmium then deposits in bulk. Isotherms with discontinuities of the type predicted by Fowler's theory have also been found for hydrogen on palladium, but here the hydrogen actually penetrates the metallic lattice itself, so that the problem is no longer a surface one, but is, strictly speaking, a problem in the theory of solutions.

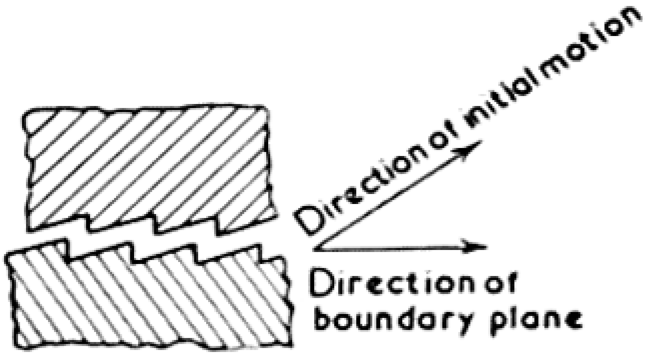


FIG. 103.

11. Friction and Lubrication

The first serious attempt to account for the phenomenon of *friction* seems to have been due to **Coulomb,** who suggested that it was due to the interlocking of surface asperities in a manner similar to the meshing of teeth of gear wheels (Fig. 103). Such an explanation certainly accounts for the main observed facts, such as the law that the force of friction is proportional to the force pressing the surfaces together, and also that it is practically independent of the area of contact. Both of these results can be understood if we merely suppose that, in order to disengage the "teeth" from one another, it is necessary to lift up the upper body slightly as well as moving it sideways. The work done in lifting the body up the inclined planes would be lost as heat when it again falls into a position with the "teeth" closely interlocked. This idea seems to have been accepted for nearly a century, and has been superseded only comparatively recently, when information on the behaviour of solids under carefully controlled conditions became available, due to the work of **Hardy** and, more recently, of **Bowden** and his collaborators.

The explanation now accepted runs on somewhat different lines. It is known from interferometric and other evidence that the surface of a crystal is never molecularly smooth, even if the most careful polishing methods are used, so that two solids placed one above the other really only "touch" at a comparatively small number of points. At these points of contact the pressure is very high, and the crystals will yield, elastically and plastically, until the areas of contact have increased and the pressure has fallen to a value at which yielding ceases. The total area of contact is then nearly proportional to the total force pressing the surfaces together and, at least until this force becomes very large, is nearly independent of the total area of the crystal faces. The effect of increasing the force pressing the surfaces together is merely to enlarge the areas of contact and possibly to produce others (Fig. 104). To account for the observed fact that friction is proportional to normal reaction it is necessary to assume that the adhesive forces are proportional to the area of contact. Now we have already mentioned (p. 267) that metals under "clean" conditions adhere to one another so tenaciously that tearing them apart results in pieces being dragged out of one surface by the other, so it is very reasonable that the total adhesion between the "peaks" in Fig. 104 should be proportional simply to the area of contact, the force pressing the surfaces together only affecting these forces by determining the area of contact.

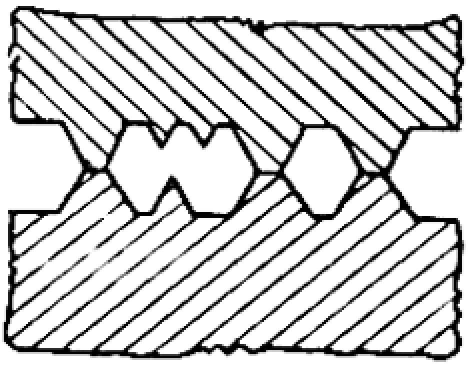


FIG. 104.

The problem of lessening friction is therefore that of keeping the peaks of the solids out of contact as far as possible, and of lessening the adhesion between them if contact is inevitable. This is done by *lubrication*, and we must consider two quite distinct problems, namely, the interaction of the oil or other lubricant with the surfaces of metals, and the design of bearings that are run in conditions where the surfaces are flooded with lubricant. Recently, the situation of "extreme pressure", in which conditions are so severe that the metal reacts chemically with the lubricant, has also been studied.

The second question is really a problem in hydrodynamics, which was first considered by **Reynolds**. He showed that, if bearings are properly designed, it can be arranged that the solid surfaces can, while in relative motion, be kept out of contact with one another by films of lubricant of appreciable thickness. The basic principle of design of such bearings is that the moving surfaces must not be exactly parallel but, instead, slightly inclined towards one another. In such circumstances, Reynolds predicted, and the prediction has been verified in practice, that the fluid films should, when sheared rapidly, be capable of supporting the heavy pressures associated with a shaft "floating" in a journal bearing.

An entirely different set of problems arises when we consider the mutual sliding of two solid surfaces with only a trace of lubricant between them. The peaks of the solid surfaces can be prevented from seizing up in two main ways. We can either interpose bodies such as graphite, which can be sheared extremely easily, or we may form adsorbed films with bodies such as soaps or fatty oils such as castor oil. As we have seen, such molecules have chemically active ends, which can "anchor" themselves firmly to other surfaces, while the paraffin chains will stand out nearly normally. Thus, instead of having to reckon with the cohesion between the bare solids, we have instead that between the chemically inactive paraffin chains. On the other hand, a purely inert oil, consisting of a mixture of paraffins, is unable to "anchor" itself to the exposed peaks of the solid surface in this way, and there is correspondingly greater likelihood of bare peaks coming together. In this way it is possible to understand that castor oil is often a more efficient lubricant than a mineral oil of comparable viscosity. Under severe conditions, the best lubricant may be a solid whose crystals shear easily.

Yet another requirement commonly arises, that of keeping the surface cool, for example, cutting tools on a lathe or the pistons of a car engine. High temperatures are undesirable for many reasons, "seizing-up" is favoured, the temper of steel may be destroyed, and the lubricant decomposed chemically. It will be seen that the choice of a lubricant is quite a difficult problem, and an "oil" or "grease" that is designed for a particular function may be a careful blend of several constituents.

BIBLIOGRAPHY FOR CHAPTER IX

N. K. Adam, *Physics and Chemistry of Surfaces*. O.U.P. (1938).

Van der Waals, *Zeit fur. Phys. Chem.*, **13**, 657 (1894).

R. H. Fowler and E. A. Guggenheim, *Statistical Thermodynamics*, Chapter X. C.U.P. (1939).

Lord Rayleigh, *Scientific Papers*, Vol. III.

F. P. Bowden and D. Tabor, *Friction and Lubrication of Solids*. O.U.P. (1950).

V. K. La Mer and R. Gruen. *Trans. Faraday Soc.* **48**, 410 (1952).

Examples on Chapter IX

1. Obtain an estimate of the radius of a drop of water such that its surface energy is equal to the heat required to convert it to vapour. How is this result modified if the drop is formed on an ion with one electronic charge? Comment on the significance of your result in determining the behaviour of a Wilson cloud chamber.

(Surface energy of water 100 ergs/cm.^{2}. Total heat required to vaporise water at room temperature, 600 cal./g. Electronic charge, $4{\cdot}8 \times 10^{-10}$ electrostatic units.)

(Hint.—The electrostatic energy of a sphere of radius r cm. with a charge of e electrostatic units is $\frac{1}{2} e^2/r$.)

2. Estimate the rise of temperature if a soap-film of mean thickness 10^{-3} cm. and specific heat 1 calorie per gram per degree is allowed to collapse. (Surface tension of soap solution 25 dynes/cm.)

3. Explain how it is possible for a soap-bubble to be stable although the thickness of its wall varies from place to place.

4. A soap-bubble is blown on the end of a capillary tube and is then allowed to contract. Assuming that the rate at which unit volume of air escapes is proportional to the excess pressure, derive a relation between the radius of the bubble and the time. Find also how the radius varies if the free end of the capillary is connected to a large bottle in which the pressure is maintained constant.

5. A film, whose thickness may be taken as uniform, is formed on a horizontal circular ring. Set up the differential equation determining its profile taking the effect of gravity into account, assuming that the deflection is everywhere small, and hence show that the profile is a parabola.

(Hint.—Consider the equilibrium of a circular portion of the film under the action of its weight and the surface tension acting round its edge.)

CHAPTER X

DIMENSIONAL METHODS. DYNAMIC SIMILARITY

1. Introduction

Dimensional analysis amounts to the systematic exploitation of a certain common-sense principle of physics, which limits very severely the possible forms of an acceptable physical statement. The principle is, without any doubt, correct, and it often enables a great deal of information to be gained about a phenomenon for which no quantitative theory exists. (An outstanding example of the power of dimensional analysis is the treatment of thermal convection. See Roberts, *Heat and Thermodynamics*.) However, the very simplicity of the reasoning seems to excite a sort of unconscious suspicion, and the whole method is not infrequently the subject of utterly barren controversy. In the author's opinion, much of the misunderstanding has arisen from a failure to separate the process of dimensional analysis into two quite distinct stages, which are related to one another in the same way as are the premises and conclusion of a logical argument.

2. Description of the process

The initial step is to decide what are the main physical magnitudes that are likely to be connected by a theoretical formula with the particular physical magnitude that we are interested in. This can *never* be decided rigorously, because any physical quantity is obviously affected to a minor extent by countless others, and it is necessary to confine oneself to those of major importance. (Dimensional analysis ceases to give results of any real value when more than about six variables are involved.) The process of deciding which variables are likely to be relevant differs for each problem, and cannot be precisely defined. It may involve, *e.g.*

(*a*) "Order of magnitude" calculations to show that the influence of certain quantities is negligible.

(*b*) Argument by analogy from other theories.

(*c*) Showing that two quantities, *e.g.* two lengths, can enter the final result only as *e.g.* a product or ratio.

(*d*) Physical insight and "horse-sense".

Once the premiss has been made, the conclusion follows as a rigorous consequence of the fact that the only acceptable format for a physical statement is one *that remains true whatever the units employed.* For example, if we chose to substitute solar for sidereal days in our measurements of time, both sides of any permissible physical equation must either be unaffected, or else must both be multiplied by the same factor. Similarly, our units of length and mass are entirely arbitrary, and changes in them cannot affect the physical content of an equation. These facts enable us to determine *all* possible types of formula that are logically consistent with a given physical premiss. (It sometimes happens that no such formula can be found, in which case the premiss must be wrong.) We use the conventional notation to exhibit in what way changes in the units affect the numerical values of our various physical magnitudes. Thus we write

$$\begin{aligned}
&\text{Velocity} &&= \text{Length per unit time} &&= [L]\,[T]^{-1}\\
&\text{Momentum} &&= \text{Mass} \times \text{Velocity} &&= [M]\,[L]\,[T]^{-1}\\
&\text{Kinetic Energy} &&= \text{Mass} \times (\text{Velocity})^2 &&= [M]\,[L]^2[T]^{-2}\\
&\text{Force} &&= \text{Mass} \times \text{Acceleration} &&= [M]\,[L]\,[T]^{-2}, \text{ etc.}
\end{aligned}$$

Since equations of this type are only meant to exhibit the indices of M, L, and T, it is unnecessary to introduce multiplying constants. To illustrate the process, we shall now work through three simple cases, giving the reasoning in full.

(*a*) *Period of simple pendulum T.*—Possible variables, m mass of bob, l length, g acceleration of gravity.

Premiss.—$T = f(m, l, g)$, where f is an unknown function. Any function that occurs in physics can be expressed as a power series for at least part of its range, so we write $f = \Sigma m^{\alpha} l^{\beta} g^{\gamma}$ where the summation is to be over all values of α, β, and γ that give the correct dimensions.

Comparison of indices.—For *each term* of f we must have

$$[T] = [M]^{\alpha}\,[L]^{\beta}\,([L]\,[T]^{-2})^{\gamma}$$

This is satisfied *only* if $\alpha = 0$, $\beta = \frac{1}{2}$, $\gamma = -\frac{1}{2}$, so that the function f reduces to a single term, because no other set of values of α, β, γ is consistent with the premiss.

Conclusion.—The only possible formula *consistent with the premiss* is

$$T \propto \sqrt{\frac{l}{g}} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (1)$$

(*b*) *Period of pendulum with large spherical bob.*—Possible variables m, l, g, and r (radius of bob).

Premiss.— $T = f(m, l, g, r)$.

Write $f = \Sigma m^{\alpha} l^{\beta} g^{\gamma} r^{\delta}$.

Comparison of indices.—For *each term* we must have

$$[T] = [M]^{\alpha} [L]^{\beta} [L]^{\gamma} [T]^{-2\gamma} [L]^{\delta}.$$

This is satisfied if $\alpha = 0, \gamma = -\tfrac{1}{2}, \beta + \delta = \tfrac{1}{2}$.

Any value of δ is therefore possible if the appropriate value of β is taken with it. This means that dimensional considerations do not suffice to fix the possible type of formula uniquely. We might have *e.g.* $T = A\left(\frac{l}{g}\right)^{\frac{1}{2}} + B\left(\frac{l}{g}\right)^{\frac{1}{2}}\left(\frac{r}{l}\right)$. Each of these terms is of the correct dimensions, so also is the sum of these two terms. The most general result consistent with the premiss is a sum of *all* possible types of term that are dimensionally correct, that is, it involves an unknown function. The function f therefore takes the form

$$\sum_{\delta} l^{\frac{1}{2}-\delta} g^{-\frac{1}{2}} r^{\delta} = \sqrt{\frac{l}{g}} \sum_{\delta} c_{\delta}\left(\frac{r}{l}\right)^{\delta}$$

where the c_{δ}'s are constants that cannot be determined by dimensional considerations. We may write this sum as

$$T = \sqrt{\frac{l}{g}}\,\phi\left(\frac{r}{l}\right)$$

which is equivalent to the assumption of an arbitrary function ϕ of r/l, whose form is left undetermined by dimensional considerations.

Conclusion.—The most general possible formula consistent with the premiss is

$$T = \sqrt{\frac{l}{g}}\,\phi\left(\frac{r}{l}\right) \quad \dots\dots\dots\dots\dots\dots \quad (2)$$

where ϕ is an arbitrary function.

(c) *Force F on a sphere of radius a, moving steadily through a liquid with velocity v.*—Possible variables a, v, ρ (density of liquid), η (viscosity of liquid).

Premiss.— $F = f(a, v, \rho, \eta) = \Sigma a^{\alpha} v^{\beta} \rho^{\gamma} \eta^{\delta}$.

Comparison of indices.—For *each term* we must have

$$[M][L][T]^{-2} = [L]^{\alpha} [L]^{\beta} [T]^{-\beta} [M]^{\gamma} [L]^{-3\gamma} [M]^{\delta} [L]^{-\delta} [T]^{-\delta}.$$

One of the four indices may be chosen arbitrarily, but the others are then fixed. We have $\alpha = \beta = 2 - \delta$, $\gamma = 1 - \delta$. The function f therefore takes the form

$$F = \sum_{\delta} c_{\delta} \rho a^2 v^2 \left(\frac{\eta}{\rho a v}\right)^{\delta}$$

$$= \rho a^2 v^2\, \phi\left(\frac{\eta}{\rho a v}\right)$$

where ϕ is an arbitrary function.

Conclusion.—The most general possible formula consistent with the premiss is

$$F = \rho a^2 v^2 \phi\left(\frac{\eta}{\rho a v}\right) \dots\dots\dots\dots\dots\dots \quad (3)$$

where ϕ is arbitrary.

The discussion of these three cases illustrates the various types of conclusion that can be reached by a dimensional argument. If not all the relevant variables are included in the premiss the conclusion will often (but not always) be that the indices cannot be chosen in a manner consistent with the above principles, so that *no* physical formula of the kind postulated is possible. A second type of conclusion is illustrated by case (*a*). Here there is only *one* type of physical formula that is consistent with the premiss. The third type of conclusion is illustrated by cases (*b*) and (*c*). Here the nature of possible formulae is determined except for an arbitrary function. The argument of such a function will invariably be a *dimensionless constant*, that is to say, a physical quantity whose value is independent of the units employed. Such a quantity may be simply the ratio of two *like* quantities, such as two lengths or two velocities, as in case (*b*), or it may be built up as a product of several dimensional quantities as in case (*c*). It can be verified that the quantity $\rho a v/\eta$, usually known as the Reynolds number, is indeed dimensionless.

The principle that we are invoking can be expressed in a slightly different way as follows: Divide both sides of equation (3) by $\rho a^2 v^2$ so that it becomes

$$\frac{F}{\rho a^2 v^2} = \phi\left(\frac{\eta}{\rho a v}\right)$$

which can also be written

$$\psi\left(\frac{F}{\rho a^2 v^2}, \quad \frac{\eta}{\rho a v}\right) = 0 \dots\dots\dots\dots\dots\dots \quad (4)$$

where ψ is an arbitrary function of the two variables, both of which are, as is easily verified, dimensionless, so that our physical relation has become an arbitrary *functional relationship between two dimensionless variables.* Thus, our process of dimensional analysis can be described in a concise, but rather formal, manner thus:—

(*a*) Decide what variables are likely to enter the physical relation sought. Write down the dimensions of each.

(*b*) Write down all the dimensionless quantities that can be formed by multiplying powers (not necessarily positive or integral) of these variables. (The process of "comparison of indices" that we have used above ensures that this is done in a systematic way, so that no such quantities are missed.)

(c) The proposed physical equation must then be expressible as a functional relationship involving one or more of these dimensionless quantities (through which alone dimensional quantities can enter the equation).

3. Model experiments. The concept of dynamic similarity

It is a well-known fact that a scale model of a machine very often does not work in the same way as the machine itself. As an obvious example of the type of difficulty that constantly arises, consider what happens if all the linear dimensions of a framework or girder are increased n times. Then all the volumes, and therefore weights, are increased by a factor n^3, but the cross-sectional areas are only increased by a factor n^2, so that the stresses (forces per unit area) are increased by a factor n on scaling up. Since yield stresses depend mainly on the nature of the material, it is quite possible that the model has a considerable factor of safety, and yet the stresses in the full-sized structure may be dangerously high. This kind of consideration is a potent factor in limiting the possible sizes of animals (particularly birds and insects, because the energy expended in flying increases much more rapidly than does the volume of the animal as the dimensions are increased).

One possible way of dealing with the problem of designing a framework from a model would be to choose deliberately a material for the model whose elastic moduli and yield stress were approximately $1/n$th of those of the material used in the actual framework itself. A second, less practical, possibility would be to use for the model a material n times as dense, but with the same elastic constants. A designer's work consists, first, of deciding what factors are likely to be most important in the design, secondly, of building the model in such a way that it gives as true a representation as possible of the effect of these factors, even if this implies a departure from strict scaling down. These two stages are closely parallel to the process of dimensional analysis that we have already outlined—we first decide which physical parameters are the most important in influencing the working of the machine, then we form these into dimensionless quantities by the process of comparing indices already described, and finally we try to arrange matters so that as many as possible of these dimensionless quantities are the same in the model as they are in the actual device. If this can be achieved for all the relevant dimensionless quantities, then we have the condition of *dynamic similarity*, and only then can we expect the behaviour of the model to be a small-scale reproduction of that of the machine. Dynamical similarity can sometimes be achieved quite easily, but when several parameters are involved considerable ingenuity may be required, or the achievement may even be theoretically impossible.

These considerations can be illustrated by the problems involved in model experiments on the resistance of ships. The following are the parameters likely to be involved.

The *linear dimensions* of the model.—The ratios of all the linear dimensions are non-dimensional quantities, but they can all be reproduced by building the model to scale and loading it so that the depth of the keel below the surface is also to scale. Let a be one of these linear dimensions (say the length of the ship or model).

The *density* ρ, and the *viscosity* η, of the water or other liquid used.

The *velocity* v, with which the model or ship travels.

The *acceleration of gravity* g. (This enters because, in addition to the resistance associated with the motion of a body through an unlimited fluid, one has also to consider the formation of *waves* at the free surface.)

The *force* F pushing the model or ship along.

The discussion of case (*c*), p. 278, shows us that two of the non-dimensional quantities that can be formed from these six physical variables are $F/\rho a^2 v^2$, which we may call the reduced force, and the Reynolds number $R = \rho a v/\eta$. A third non-dimensional number is v^2/ga, sometimes known as Froude's number. The reader can satisfy himself by a little experimenting that no other non-dimensional quantity can be built up from the six physical variables that we have mentioned, or, more precisely, that *any other* non-dimensional quantity depending only on these variables can be expressed in terms of the three that we have already found.

Let us suppose that the linear dimensions of the model are to be $1/n$th of those of the ship. It at once becomes obvious that *perfect* dynamical similarity cannot in practice be attained. Let us consider the possibilities associated with the five other physical variables enumerated, the scale of the model having been decided.

g cannot be varied significantly without travelling to another planet.

ρ cannot be varied to any great extent. Very few liquids, apart from mercury, have densities as great as 2 or as small as $\frac{1}{2}$ g./cm.^{3}.

η can be increased to any desired extent by using more viscous liquids, and could be reduced by a small factor by heating the water or by using a more mobile liquid.

v and F can be varied within very wide limits.

By choosing v for the model to be $n^{-\frac{1}{2}}$ of its value for the ship we can make Froude's number the same for both. If we then choose F for the model to be n^{-3} of its value for the ship, the reduced force is the same for both. To make the Reynolds number the same for both the ratio η/ρ for the model would have to be $n^{-\frac{3}{2}}$ of its value for full scale. If n is of the order of 100, this is obviously quite impracticable. It would, however, be possible to make Reynolds number the same for both if we were prepared to accept a discrepancy in Froude's number. For this the model would have to travel n times as fast as the full-scale ship, η being constant. We thus see that, even in this comparatively simple case, true dynamical similarity is not, in practice, attainable. To get over this difficulty, the resistance of ships is divided into "wave-making resistance" (controlled almost entirely by Froude's number) and "skin resistance" (controlled by Reynolds number). Instead of assuming an absolutely arbitrary function the assumption is made that these two contributions to the resistance can simply be added together. The most general possible result is $\frac{F}{\rho a^2 v^2} = \phi\left(\frac{\rho a v}{\eta}, \frac{v^2}{ga}\right)$ where ϕ is an arbitrary function. Instead of this we assume:

$$\frac{F}{\rho a^2 v^2} = f_1\left(\frac{\rho a v}{\eta}\right) + f_2\left(\frac{v^2}{ga}\right).$$

If this is correct it may be possible to determine f_1 and f_2 separately from the model experiments and then scale them up separately. If, for example, one was investigating a region of velocities for which the skin resistance was shown to be negligible, one could plot the graph of reduced force against Froude's number from the model experiments, and then use this graph for the full-scale ship with confidence. This situation would be one of approximate dynamical similarity.

It will be seen that these dimensional considerations are of very great help to physicists and engineers, and they are, in fact, indispensable in dealing with such questions as the design of wind-tunnels and in determining the proper velocities and pressures of air needed to reproduce a given set of conditions adequately on the model scale. They are, however, no substitute for physical insight. For example, in the question of resistance of ships already discussed, there is no *a priori* reason why the *surface tension* should not be a significant parameter. (We know, for example, that it is significant in the formation by the wind of waves at sea.) If it were significant in this context, we should have yet another similarity relation to satisfy, and we cannot satisfy those we already have. However, an "order of magnitude" calculation quickly disposes of the idea that surface tension can have any direct bearing on the resistance of a ship.

4. The checking of physical equations. Special systems of units

The reader will probably be familiar with the use of dimensional methods in checking the development of a physical theory as it proceeds. This is distinct from the applications we have just been discussing, in which we develop the *most general possible form* of a physical relation consistent with certain prior assumptions. We consider now the development of a *particular theory* from the initial assumptions to the final formula that is to be compared with experiment. It must be verified that the initial assumptions, the intermediate steps, and the final conclusion are all dimensionally correct. This serves not only the mundane purpose of checking the algebra, but helps one to appreciate the physical significance of the theory. Again, it is sometimes of assistance in the analysis of experimental results. It often happens that, within the accuracy of the experiment, two or more functions of entirely different mathematical form will represent the data equally well. If, for example, we were proposing to express the relation between two physical quantities X and Y by means of some relation like $X = AY^n$, we could determine the dimensions of the quantity A, and ask ourselves whether it was likely to be physically significant. If no interpretation of it suggested itself, it might be advisable to look for some other form of equation. However, it does sometimes happen that a physical quantity has dimensions whose interpretation is not obvious. Consider, for example, the gravitational constant G. From Newton's inverse-square law of force, the reader will easily verify that G has dimensions $[M]^{-1}[L]^3[T]^{-2}$ although no obvious interpretation of this fact suggests itself, in spite of the fundamental nature of the law in question.

Relationships such as the one we have just mentioned often occur in physics, for example, we write the law of uniformly accelerated motion in the form: $S = u_0t + \frac{1}{2}ft^2$, whereas, to express it as a relation between *dimensionless* quantities, one would have to write, *e.g.*, $S/u_0t = 1 + ft/2u_0$. The coefficients of t and t^2 appearing in the first presentation of this equation are not independent of the units employed and are consequently often known as *dimensional constants*, but the constants (1 and $\frac{1}{2}$) appearing in the second form of the equation are pure numbers. Unfortunately, certain systems of units are introduced from time to time which are so chosen that certain constantly occurring quantities, *e.g.* the gravitational constant or the velocity of light, are unity, so that they can be omitted from the physical equations in which they occur. This saves both writing and computational work, but sometimes obscures the physical meaning of an equation, and deprives us of the possibility of making dimensional checks. This can be avoided by writing the equations in their correct dimensional form, including even magnitudes that are numerically equal to unity, but, if this is done, much of the advantage of such "rational" systems disappears. Such units are admirably suited to the process of determining the numerical consequences of a definitely accepted theory, but they are apt to confuse the issue alarmingly if they are used during the process of developing a new one. These objections do not apply to the process of expressing an equation in *non-dimensional* form, *i.e.* as a relation between two or more dimensionless quantities, but only to the use of units which lead to *dimensional constants* which have a numerical value of unity. The dangers of such a proceeding should be obvious, but they have been emphasised because they have led distinguished workers into committing errors.

BIBLIOGRAPHY FOR CHAPTER X

A. W. Porter, *The Method of Dimensions.* Methuen Monograph (1933).

N. Campbell, *Measurement and Calculation.* Longmans (1928).

J. K. Roberts, *Heat and Thermodynamics.* Blackie, Glasgow (1951).

Examples on Chapter X

1. On the assumption that the velocity of waves on the surface of water depends only on the depth of water, the wavelength, the acceleration of gravity, the density and the surface tension, determine the most general possible formula connecting these quantities.

2. Surface tension can be determined by timing the period of the oscillations of a small drop of liquid about its equilibrium spherical form. Suggest a formula.

3. Suggest a formula for the variation of velocity of propagation of sound of finite amplitude in a gas, on the assumption that it depends only on the amplitude, density, pressure, and wavelength.

4. Determine a formula for the relationship between the maximum bending moment G in a beam of square section, the total load W, the unsupported length l, the side of the square section a, and Young's modulus Y, on the assumptions that no other modulus of elasticity is effectively concerned, and that W and l only enter the result through the product Wl.

5. Suggest a formula for the maximum possible height of a flagstaff, considered to be of uniform radius a, that will not buckle under its own weight, assuming that the other variables concerned are Young's modulus Y and the density ρ. Why is it reasonable that the ultimate breaking stress of the material should be absent from the formula?

6. Derive a formula giving a relationship between the mass of a planet, the gravitational constant, the period of the orbit of a satellite, and the mean radius of that orbit. Why, and in what circumstances, is it correct to omit the mass of the satellite from the formula?

MATHEMATICAL APPENDIX

We derive here a number of results required in the text. They have been collected in one place, both because some of the derivations are rather long, and also because they are used many times in this book and all have a large number of applications in various other branches of physics not treated in this book.

1. The damped simple harmonic oscillator

To fix our ideas, consider a particle of mass m, attached by a spring to a fixed point, which we may take as the origin, and suppose that the force on it is mp^2x directed towards the origin, where x is the displacement from the equilibrium position. Let it be subject also to a frictional force proportional to its velocity, which will always tend to decrease this velocity. Let this force be $k\,dx/dt$. Then, by Newton's second law, the equation of motion is

$$m\frac{d^2x}{dt^2} = -k\frac{dx}{dt} - mp^2x + F(t) \quad \ldots\ldots\ldots\ldots\ldots \quad (1)$$

where the first two terms on the right-hand side have already been described, and $F(t)$ represents any other force acting on the particle due to external causes.

(*a*) *The case of no external applied force. Free oscillations.*—If $F(t) = 0$, equation (1) becomes

$$\frac{d^2x}{dt^2} + \frac{k}{m}\frac{dx}{dt} + p^2x = 0 \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (2)$$

This is a differential equation for x which is linear, *i.e.* it is a linear function of x and its differential coefficients. Such equations are considered in textbooks on differential equations. They have the important property that, if two functions $x_1(t)$, $x_2(t)$ satisfy such an equation separately, then it will also be satisfied by $Ax_1(t) + Bx_2(t)$, where A and B are constants independent of t and x, as can easily be verified by substitution. Also, the general theory of differential equations shows that an equation involving no differential coefficients of x beyond the *second* has, as its most general possible solution, an expression containing exactly *two* quantities such as A and B that can be chosen at will. Equation (2) can therefore be solved completely if we can determine two *distinct* solutions, $x_1(t)$ and $x_2(t)$. This can always be done if, as in equation (2), the coefficients of x, $\frac{dx}{dt}$ and $\frac{d^2x}{dt^2}$, are independent of t. To do this, assume as a trial solution $x = e^{qt}$. Substituting in equation (2), we see that it will be satisfied if

$$q^2 + \frac{kq}{m} + p^2 = 0,$$

which will be true for the following two values of q:

$$\begin{matrix} q_1 \\ q_2 \end{matrix} = -\frac{k}{2m} \pm \sqrt{\left(\frac{k^2}{4m^2} - p^2\right)} \quad \ldots\ldots\ldots\ldots\ldots \quad (3)$$

We thus see that we have two distinct solutions of (2), except in the special case $p = k/2m$, which we will consider in a moment. It is convenient to consider separately the cases $k > 2mp$ and $k < 2mp$.

(i) $k > 2mp$. *The over-damped oscillator.*—In this case, both q_1 and q_2 are real. Since k is an essentially positive quantity q_2 is negative q_1 is also negative because $(k^2/4m^2 - p^2)^{\frac{1}{2}}$ is numerically less than $k/2m$. The complete solution is therefore

$$x = Ae^{q_1 t} + Be^{q_2 t} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (4)$$

This corresponds to the particle "creeping back" exponentially to its equilibrium position after being displaced. A and B are not determined by the equation of motion itself, but can be found, *e.g.* if the position and velocity of the particle are known at some definite time (say $t = 0$). For then we have $x = A + B$, $dx/dt = Aq_1 + Bq_2$, which enables A and B to be determined. If the particle is just displaced and released, so that $dx/dt = 0$, x, the displacement, approaches the origin asymptotically with time from one side. If the particle is displaced and then projected towards the origin, it may pass through the origin and overshoot it once, but will then approach it asymptotically from the other side.

(ii) $k < 2mp$.—In this case q_1 and q_2 are complex conjugate quantities. (4) still satisfies (2) formally, but we require a solution in real form for physical applications. Solution (4) is real if A and B are also complex conjugate quantities, which we may write as $\frac{1}{2}(C \pm iD)$ where C and D are real. Let us write $n^2 = p^2 - k^2/4m^2$, so that equation (3) becomes $\begin{matrix} q_1 \\ q_2 \end{matrix} = \pm in - k/2m$. Substituting these values in (4) we get, simplifying and using Demoivre's theorem,

$$x = e^{-kt/2m}(C \cos nt + D \sin nt) \quad \ldots\ldots\ldots\ldots \quad (5)$$

It can be verified by substitution that this satisfies (2) for any values of C and D and is therefore an alternative form of the general solution.

This solution corresponds to a *damped* simple harmonic oscillation, representing sinusoidal oscillations which decay with time because of the factor $e^{-kt/2m}$. The quantity in the brackets will repeat itself at intervals of time corresponding to increases of nt by 2π, 4π, etc. In particular, it will attain the value zero *twice* in each such interval, so that, if the particle is displaced and then released, it will pass through the origin infinitely many times before finally coming to rest there. As in case (i), C and D can be determined from a knowledge of x and dx/dt at any instant.

It can be verified by substitution that the particle will pass through the origin whenever t is given by $t = t_1 + r\pi/n$, where t_1 is the time of the first transit, *i.e.* the smallest solution of the equation $C \cos nt + D \sin nt = 0$, and r is a positive integer. Its periodic time is thus

$$T = 2\pi/n = \frac{2\pi}{p}\left(1 - \frac{k^2}{4m^2p^2}\right)^{-\frac{1}{2}},$$

which, if k is small compared with mp, can be written

$$T = \frac{2\pi}{p}\left(1 + \frac{k^2}{8m^2p^2} + \dots\right) \qquad (6)$$

Now $2\pi/p$ is the periodic time that the oscillations of the particle would have if there were no damping, and it will thus be seen that the effect of the damping is not only a progressive reduction of amplitude, but also a slight *lengthening* of the periodic time. These two effects are connected by equation (6).

It is customary to express the decay of oscillations by means of the *logarithmic decrement*, which is the logarithm (to base e) of the ratio of *two successive maxima on opposite sides of the origin*. We have already seen that successive *zeros* of x are spaced by time intervals of π/n, and, provided that k/mp is small compared with unity, successive *maxima* of x are also spaced by this interval. This statement is not quite true, because the turning-points of expression (5) are not precisely the same as the maxima of the expression ($C \cos nt + D \sin nt$), but it is a sufficiently good approximation as long as k/mp is small. During the time π/n, the amplitude of the oscillation decays by the factor $\exp(k\pi/2mn)$ from (5), or, if δ is the logarithmic decrement, we may write

$$\delta = \frac{k\pi}{2mn} \qquad (7)$$

Very often, *e.g.* in a pendulum experiment, we require the "ideal" period $2\pi/p$ that we should get if there were no damping. The process of determining it is then—

A. To determine the actual period $2\pi/n$.

B. To determine δ by a comparison of successive maxima, or, if the damping is slight, by measuring the decrease of amplitude after say, 10 or 100 swings.

C. The correcting factor is then $1 + k^2/8m^2p^2 \simeq 1 + \delta^2/2\pi^2$ from (7).

(iii) *The case* $k = 2mp$ (*"critical damping"*).—In this case $q_1 = q_2$ and we only get one solution of equation (2). The complete solution can, however, be found as a limiting case of equation (5) as the damping is gradually increased, for $k \to 2mp$ implies $n \to 0$. Now, as $\theta \to 0$, we know that $\cos\theta \to 1$ and $\sin\theta \to \theta$, which suggests that the limiting form of (5) is

$$x \to e^{-kt/2m}(C + Dnt).$$

The first term corresponds to our known solution, and the second term suggests that we try $x = te^{-kt/2m}$ for a second solution, and we can verify by actual substitution that it is a solution. We thus have

$$x = e^{-kt/2m}(C + D't) \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (8)$$

as our complete solution in this limiting case. The properties of this solution are similar to those of case (i) already discussed. The final result is an exponential creep towards the origin, with the possibility of one "overshoot" but not more.

A result of some interest can be deduced for this case. Let us suppose that we are in a position to vary the damping k, but not the natural period $2\pi/p$. (For example, the damping of a moving coil galvanometer depends on the external circuit, but the natural period cannot easily be altered, as it depends only on the suspension and the moment of inertia of the moving parts.) Then, if we draw our particle aside a distance x and then release it, we can define a time, say t_{100}, at and after which the particle is never more than at a distance $0{\cdot}01x$ from the origin. Then this time, for a fixed natural period $2\pi/p$, is a minimum when k corresponds to critical damping. Expressed more loosely, critical damping corresponds to the most rapid "settling down" of x to its equilibrium value. If the damping is made smaller, the equilibrium point is constantly overshot, if it is made larger, the "creep" towards equilibrium is unnecessarily slow. We omit the proof of this result, as it is lengthy, though quite straightforward. The figure 0·01 has been chosen arbitrarily—an analogous result is true for any other number < 1.

(*b*) *The effect of an external disturbance. Forced oscillations.*—Equation (1) can be solved for any variation of external force with time. The reader is referred to treatises on differential equations for descriptions of various methods of doing this. We notice first of all that, if we can determine *any* solution of equation (1), say $x = x_0(t)$, the equation will also be satisfied by $x = x_0(t) + Ae^{q_1 t} + Be^{q_2 t}$, where q_1 and q_2 are given by (3). This result is a consequence of the linearity of equation (1), that is, because x and its derivatives only appear linearly, and never as powers or products. Thus, to solve equation (1), it is sufficient to find one particular solution, sometimes known as a *particular integral*, and add to that the general solution of equation (2) in the form (4), (5), or (7) as appropriate. This solution is usually referred to as the *complementary function* of (1). The complete solution now involves two constants that can be varied independently, and is thus the most general solution possible. The only case that we shall consider in detail is that of an applied force that varies sinusoidally with time, so that we may put $F(t) = F_0 \cos \omega t$. We shall see later, when considering Fourier's theorem (p. 291), that the solution for any form of $F(t)$ likely to occur in practice could, in principle, be inferred from this one. Equation (1) may now be written

$$\frac{d^2x}{dt^2} + \frac{k}{m}\frac{dx}{dt} + p^2x = \frac{F_0}{m}\cos \omega t \quad \ldots\ldots\ldots\ldots\ldots \quad (9)$$

We look for a particular solution of (9) of the form

$$x = X_0 \cos(\omega t - \gamma) = X_0 \cos\gamma \cos \omega t + X_0 \sin\gamma \sin \omega t \quad \ldots(10)$$

Substituting this in equation (9) and equating coefficients of $\cos \omega t$, and $\sin \omega t$, we find that (10) is a possible solution of (9) provided that

$$X_0 = \frac{F_0}{m}\left[(p^2 - \omega^2) + \frac{k^2\omega^2}{m^2}\right]^{-\frac{1}{2}}, \ \tan\gamma = \frac{k\omega}{m(p^2 - \omega^2)} \quad \ldots (11)$$

The most general solution of (9) is now obtained by adding the complementary function (4), (5), or (7) to (10). Solution (10) behaves in quite a different way from the complementary function, because the latter always contains an exponential damping factor for finite k, and therefore becomes unimportant for sufficiently long times. Solution (10) thus represents the *steady* response of the particle to an oscillating force and is usually known as the *forced* oscillation. This solution is *independent of the initial conditions* at $t = 0$. The complementary function contains two arbitrary constants, which have to be determined from the initial conditions, and represents the *free* oscillations of the particle, which could occur even if there were no force $F(t)$ at all. If a sinusoidal force is *suddenly* applied at $t = 0$ we shall have to take account of both free and forced oscillations. After a time the free oscillations die away, and only the forced oscillations persist. In many problems, only the forced oscillations are of interest but in others the free oscillations have to be considered. Because of their decay with time, the terms corresponding to the *free* oscillations are often known as *transients*.

The constant X_0 in equation (11) is plotted as a function of the forcing frequency $\omega/2\pi$ in Fig. 105, together with the variation of γ. It will be seen that, as the forcing frequency increases so does γ, which means that the forced oscillations are lagging behind the force causing them. When $\omega = p$, $\tan \gamma$ is infinite so that γ is $\pi/2$. As ω is increased still further, γ increases, approaching a value of π as ω becomes very large. By differentiation it can be shown that X_0 is a maximum when

$$\omega^2 = p^2 - \frac{k^2}{2m^2} \quad \dots\dots\dots\dots\dots\dots \quad (12)$$

and that the value of X_0 is then approximately equal to F_0/kp. Thus the height of the maximum value of X_0 is *inversely proportional to the damping*. The occurrence of the maximum is called *resonance*, and the corresponding period of $F(t)$ is then nearly equal to $2\pi/p$, the period of free oscillation if there were no damping. Strictly speaking, the three quantities: period of free oscillation with damping, period of free oscillation without damping, and period of forced oscillation at resonance, are distinct from one another, but the differences between them are not large unless the system is heavily damped.

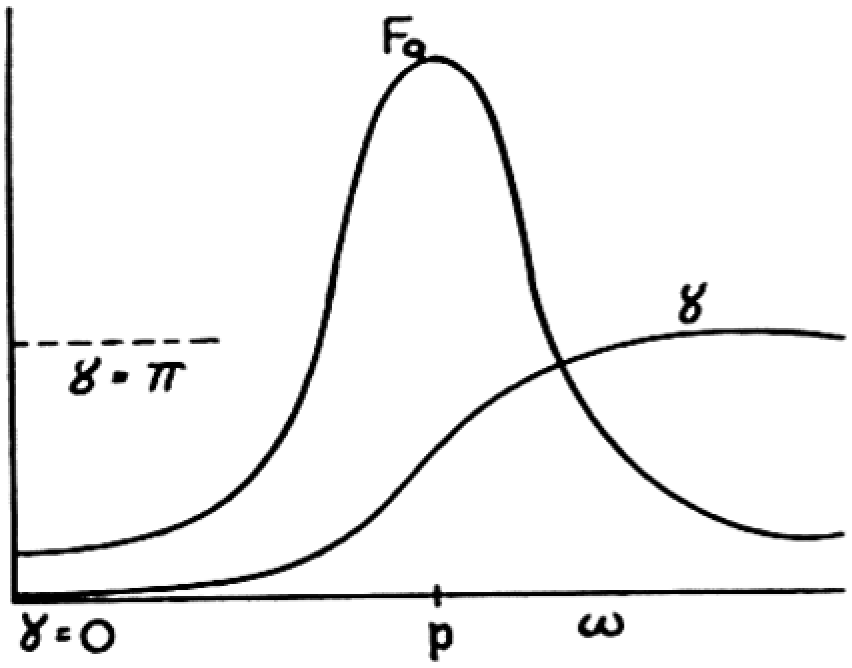


FIG. 105.

The effects of non-linearity in the oscillation equation. There are two main ways in which equation (1) may fail to describe the facts, when the oscillations become of large amplitude. The *frictional* force may no longer be proportional to the velocity, and the *restoring* or elastic force may depend on the displacement in some non-linear fashion. Equation (1) will almost always be valid at sufficiently low amplitudes of oscillation, and its solution can therefore always be used as a first approximation. The appearance of non-linear terms in the differential equation gives rise to two main types of effect. First, the resonance curve may be modified—in particular, discontinuities may appear in it and the response to a sinusoidal disturbance, $F_0 \cos \omega t$, may not be unique, but may depend on the initial conditions, or on whether we have approached ω from the direction of higher or of lower frequencies. Secondly, if two forcing oscillations of periods $2\pi/\omega_1$ and $2\pi/\omega_2$ are separately applied, in the linear case, each disturbance produces its own forced oscillation *just as if the other were not there*, but in the non-linear case, not only is the amplitude of each forced oscillation affected by that of the other, but we also get a whole set of new forced oscillations, with periods such as $2\pi/(\omega_1 - \omega_2)$, $2\pi/(\omega_1 + \omega_2)$, $2\pi/(2\omega_1 + \omega_2)$, etc., which are not present at all in the linear case, these being known sometimes as *combination tones*. The reader is referred to textbooks on sound and on valve-oscillators for a detailed discussion of these effects.

2. Fourier's theorem

This theorem states that, subject to certain conditions which are almost always fulfilled by functions of physical interest, any function

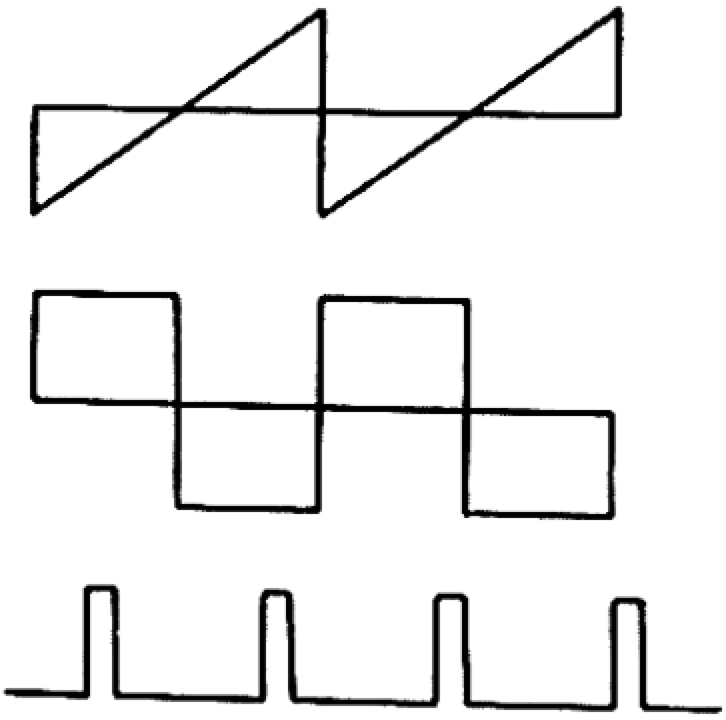


FIG. 106.

of a variable t that repeats itself in an interval T can be expressed in the form:—

$$f(t) = \tfrac{1}{2}a_0 + \sum_{r=1}^{\infty} \left\{ a_r \cos\left(\frac{2\pi r t}{T}\right) + b_r \sin\left(\frac{2\pi r t}{T}\right) \right\} \quad \ldots\ldots \quad (13)$$

The theorem applies not only to ordinary "well-behaved" functions, but to functions such as those shown in Fig. 106, which possess discontinuities and "sharp corners". Indeed, the only types of "common" function to which the theory does not apply are those like $\sin(1/x)$, which oscillates very violently over part of the range of x, or functions like tan x which have infinities. We shall now show how the coefficients a_r, b_r can be determined if $f(t)$ is given. We shall not attempt a formal proof of the possibility of an expansion such as (13), which is a difficult problem of pure mathematics, but we shall show that, *assuming* the possibility of such an expansion, the coefficients a_r, b_r can be uniquely determined. The proof depends on the fact that, if r and s are positive integers or zero, the integrals

$$\int_0^{2\pi} \cos r\theta \sin s\theta \, d\theta = 0 \quad \dots\dots\dots\dots\dots \quad (14)$$

all vanish, while the integrals

$$C_{rs} = \int_0^{2\pi} \cos r\theta \cos s\theta \, d\theta \text{ and } S_{rs} = \int_0^{2\pi} \sin r\theta \sin s\theta \, d\theta \quad \dots \quad (15)$$

vanish unless $r = s$, in which case they take the value π except that $C_{00} = 2\pi$ and $S_{00} = 0$. The relations (14) and (15) may be summed up in words by saying that, if we select two *different* functions from the set 1, $\cos\theta$, $\sin\theta$, $\cos 2\theta$, $\sin 2\theta$ . . . and multiply them together, the result of an integration from 0 to 2π is zero. This result is easily proved by expressing, *e.g.*, $\cos r\theta \cos s\theta$ as $\frac{1}{2}\cos(r+s)\,\theta + \frac{1}{2}\cos(r-s)\,\theta$, both of which terms when integrated contribute nothing to the integral C_{rs}, except in the special case when $r = s$. Sets of functions that have this property (that when two distinct members are multiplied together and integrated over a certain domain the result vanishes), are said to form an *orthogonal set*. A great many such sets of functions are known, and play a fundamental part in modern mathematical physics. The process of solving a problem by the use of orthogonal functions is usually the following:

(*a*) Select an orthogonal set of functions that individually satisfy the boundary conditions.

(*b*) Substitute such a series as (13) into the differential equation.

(*c*) Make use of the fact that nearly all of the integrals like (14) and (15) vanish.

To determine the coefficient a_r in equation (13), multiply both sides of it by $\cos(2\pi rt/T)$ and integrate from $t = 0$ to $t = T$. Then, assuming the possibility of integrating the series on the left term by term, the only term that survives on the left-hand side is

$$\int_0^T a_r \cos^2\left(\frac{2\pi rt}{T}\right) dt = \frac{a_r T}{2},$$

so that we have

$$a_r = \frac{2}{T}\int_0^T f(t) \cos\left(\frac{2\pi rt}{T}\right) dt \quad \dots\dots\dots\dots \quad (16)$$

Similarly,

$$b_r = \frac{2}{T}\int_0^T f(t)\sin\left(\frac{2\pi rt}{T}\right)dt \qquad \ldots\ldots\ldots\ldots\ldots (17)$$

If we integrate equation (13) as it stands, all the terms on the left vanish except that involving a_0, and we have

$$a_0 = \frac{2}{T}\int_0^T f(t)\,dt,$$

so that equation (16) is also true when $r = 0$. [This is the point of using $\frac{1}{2}a_0$ instead of a_0 in equation (13).] The integrals occurring in (16) and (17) are to be interpreted in the following way. If $f(t)$ can be represented by the same algebraic expression throughout the range $0 \leqslant t \leqslant T$, they are ordinary integrals. If, however, $f(t)$ is such a function as

$$f(t) = t, \quad 0 \leqslant t \leqslant \frac{T}{2}$$

$$f(t) = \frac{T}{2}, \quad \frac{T}{2} \leqslant t \leqslant T,$$

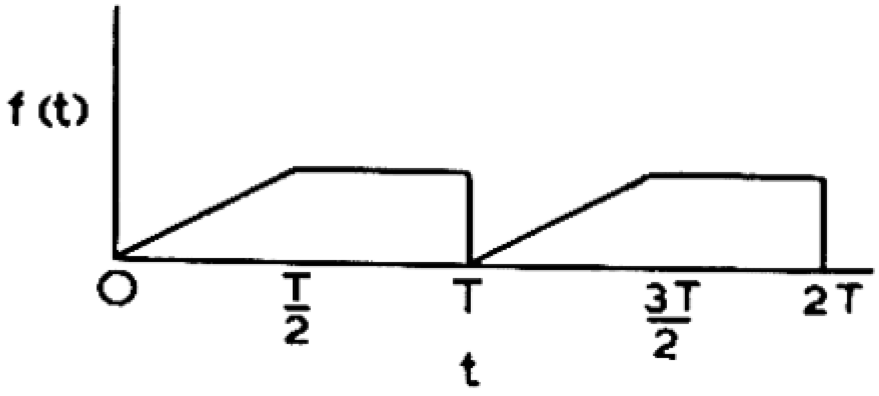


Fig. 107.

the graph of which is shown in Fig. 107, then equation (16) would have to be interpreted as

$$a_r = \frac{2}{T}\int_0^{T/2} t\cos\left(\frac{2\pi rt}{T}\right)dt + \frac{2}{T}\int_{T/2}^{T} \frac{T}{2}\cos\left(\frac{2\pi rt}{T}\right)dt,$$

the range of integration being broken at each point at which the algebraic form of $f(t)$ changes.

This process enables us to determine a_r and b_r in cases where an expansion such as (13) is possible, and, as we have mentioned, this is so for a very large class of functions, even though they may have discontinuities or "sharp corners". It is usually permissible to differentiate the series (13) term by term when it represents a "smooth" function but care may be required in the neighbourhood of a "sharp corner". It should be strongly emphasised that, if such an expansion can be made, the series (13) is not merely an *approximation* to $f(t)$, but, can be made to differ from it by as little as we please, by taking enough terms of the series, and, at an actual discontinuity, the sum of the series usually converges to the *mean* of the values on the two sides of the discontinuity.

Half-range Fourier series.—So far we have supposed $f(t)$ to have a period T, and to be known throughout one period. Let us now suppose that in the first instance we only specify the form of $f(t)$ between $t = 0$ and $t = \frac{1}{2}T$. We cannot now determine the a_r and b_r because expressions (16) and (17) assume knowledge of $f(t)$ over the whole of a period. Plainly, we can make any assumption we like about the behaviour of $f(t)$ between $t = \frac{1}{2}T$ and $t = T$. To each of these assumptions corresponds a different set of the coefficients a_r and b_r, and we thus have the possibility of having two or more trigonometrical expansions such as (13), *all of which have exactly the same form between $t = 0$ and $t = \frac{1}{2}T$*, that is, *two quite different functions can coincide over a finite range of the variable.* (Such behaviour can be proved to be impossible for two functions represented by ordinary power series expansions.) It is, however, *not* possible for two functions to coincide over the *whole* of the fundamental period T. Put in another way, if a function is known to vanish over the *whole* of the fundamental period, it can be formally proved that all its a_r and b_r must vanish, a result that is suggested by equations (16) and (17).

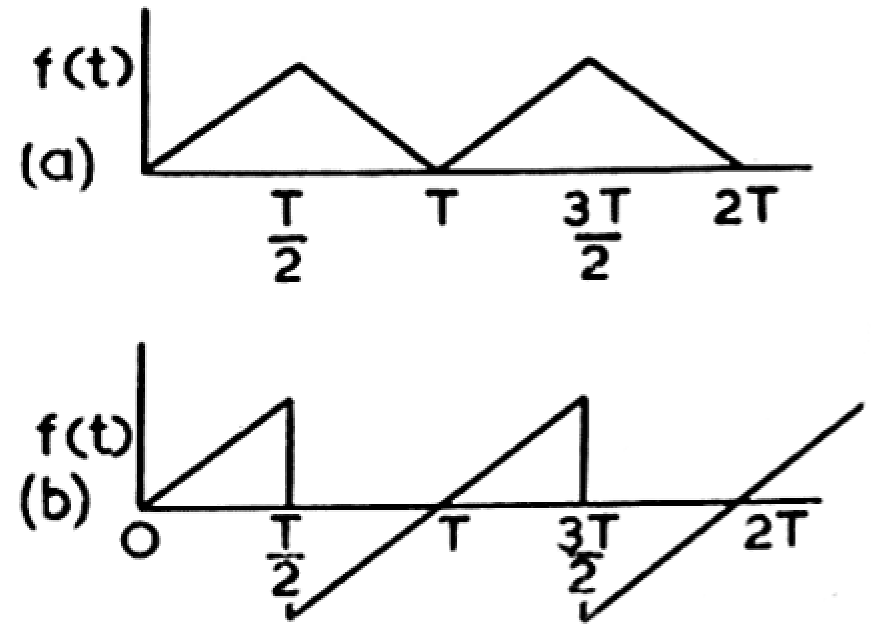


FIG. 108.

In Fig. 108 we illustrate two special ways of "completing" the function $f(t)$, that is specified to have the form $f(t) = t$ over the range $0 \leqslant t \leqslant T/2$. Fig. 108($a$) represents the function

$$f(t) = t, \quad 0 \leqslant t \leqslant T/2; \quad f(t) = T - t, \quad T/2 \leqslant t \leqslant T \quad .. (18)$$

while Fig. 108(b) represents the function

$$f(t) = t, \quad 0 \leqslant t < T/2; \quad f(t) = t - T, \quad T/2 < t \leqslant T \ ... (19)$$

In each case, the behaviour of $f(t)$ for values of t outside the range between 0 and T is specified by the requirement that $f(t + T) = f(t)$. It can easily be verified that the function represented by (18) consists entirely of a series of cosines, all the integrals for b_r vanishing, while for the function represented by (19), all the integrals for a_r vanish. Put in another way, we have "filled in" the missing half-period in (18), in such a way that the complete function is *even* in t, and is unaffected by replacing t by $-t$, which is just the property of the cosine functions, while in (19) the "filling in" has been done in such a way that $f(t)$ is an *odd* function, whose sign is changed with that of t, which is a property of all sine functions. For this reason, series such as those representing functions (18) and (19) are known as *half-range cosine* and *half-range sine* series respectively. It can easily be verified that the *two* sets of functions 1, $\cos\theta$, $\cos 2\theta$, $\cos 3\theta \ldots$, and $\sin\theta$, $\sin 2\theta$, $\sin 3\theta \ldots$, *each* form orthogonal sets if the range of integration is $\theta = 0$ to $\theta = \pi$, so that *either* set can be used to express an arbitrary function of θ over such a range. [The coefficients are determined by integrals like (16) or (17), but over the range $\theta = 0 \rightarrow \pi$.] Over this range of integration, sine and cosine functions are not orthogonal to one another.

It very often happens that one is interested only in the values of a function over a finite range of values of the variable, for example, if we are considering the vibrations of a string of length l. As we shall see below (Art. 3) it is very often helpful to express the function representing the displacement of the string as a Fourier series. The exact type of series that we choose depends on the boundary conditions at the ends of the range. For example, in discussing the oscillations of a string fixed at the points $x = 0$ and $x = l$, we should use the functions $\sin(r\pi x/l)$, each of which also satisfies these boundary conditions, that is to say we should use a *half-range sine series*. A *half-range cosine series* would be applicable, *e.g.* to the problem of finding a velocity potential ϕ (see Chapter VI, p. 156), for a fluid confined between rigid surfaces at $x = 0$ and $x = l$, the appropriate boundary conditions being $\partial\phi/\partial x = 0$ at these points. Thus, we have two distinct types of physical situation in which the use of Fourier analysis may be helpful.

(*a*) To represent a function that is known to be periodic in the variable (usually the time variable). We expand as a full-range series.

(*b*) To represent a function that is of physical interest only for a certain range of values of the variable. We expand, usually as a half-range or similar type of series whose fundamental period is some simple multiple, usually one, two, or four, of the range of the variable in question. We usually work with a series, *each term* of which satisfies the boundary conditions separately. In application (*a*) the main property of interest is the *periodic* property of sines and cosines, in application (*b*) it is the *orthogonal* property that we want, and here other sets of orthogonal functions are very often used, the precise choice depending on the boundary conditions.

3. Application of Fourier analysis to the wave equation

To illustrate these considerations we shall consider the equation of plane waves, or the one-dimensional wave equation, which is

$$\frac{\partial^2 p}{\partial x^2} = \frac{1}{c^2}\,\frac{\partial^2 p}{\partial t^2} \quad \text{(20)}$$

This has been met several times in this book, for example, it governs the propagation of sound, p. 156, and also of compressional and shear waves in a solid, pp. 107, 108. A similar equation occurs in many other physical contexts outside the scope of this book, and, in fact, seems to govern, to a first approximation, very many of the processes where a reversible conversion of energy, *e.g.* from kinetic into potential, is involved. We shall now illustrate the method by considering a simple example, explaining each step in detail. We shall suppose that p represents the excess pressure in a sound-wave, and that p is to vanish at $x = 0$ and $x = l$. This is practically the problem of sound in an air column, of length l, with *open* ends.

The steps in the argument are as follows:—

(*a*) At any instant *any* physically reasonable function satisfying these boundary conditions at $x = 0$ and $x = l$ can be expanded, *qua* function of x, as a half-range sine series. In particular this is true of p, thus we may put

$$p = \sum_r b_r \sin\left(\frac{r\pi x}{l}\right) \quad \text{(21)}$$

This is simply a mathematical fact.

(*b*) At any other instant, a similar expansion of p can be performed, but the b_r's will now have changed. That is, the variation of p as a function of time and distance can be fully described if, in equation (21), we allow the b_r's to be functions of the time and not simply constants. [Any variation of p with x is fully looked after at each instant, by the functions $\sin(r\pi x/l)$, so that it is not necessary to allow the b_r's to vary with x as well as with t.]

(*c*) So far we have assumed only that p is some physically reasonable function of the variables x and t satisfying the boundary conditions, and we have not used equation (20). Substituting expression (21) into equation (20), (this may need justification because it implies term-by-term differentiation of an infinite series), we get, remembering that b_r is independent of x,

$$0 = \frac{\partial^2 p}{\partial x^2} - \frac{1}{c^2}\,\frac{\partial^2 p}{\partial t^2} = \sum_r \left(-\frac{r^2\pi^2}{l^2}\,b_r - \frac{1}{c^2}\,\frac{d^2 b_r}{dt^2}\right)\sin\left(\frac{r\pi x}{l}\right) \;.\;. \quad (22)$$

This equation has to hold for all values of t, and for all values of x between 0 and l, which means that the right-hand side of equation (22) is a *Fourier series representing a null-function*, and the general theory of such series shows that this is possible only if each of the coefficients of sin $(r\pi x/l)$ vanishes separately. This means

$$\frac{d^2 b_r}{dt^2} + \frac{r^2\pi^2 c^2}{l^2}\,b_r = 0 \;\ldots\ldots\ldots\ldots\ldots\ldots \quad (23)$$

The general solution of this equation is

$$b_r = B_r \cos\left(\frac{r\pi ct}{l}\right) + C_r \sin\left(\frac{r\pi ct}{l}\right)$$

where B_r and C_r are arbitrary constants. The full solution of (20), the most general one that satisfies these boundary conditions, is therefore

$$p = \sum_r \left[B_r \cos\left(\frac{r\pi ct}{l}\right) + C_r \sin\left(\frac{r\pi ct}{l}\right)\right]\sin\left(\frac{r\pi x}{l}\right) \;\ldots\; \quad (24)$$

and the problem is reduced to that of determining the constants B_r and C_r. These cannot be found from the differential equation, but must be determined from the initial conditions or other data. Expression (24) represents a superposition of waves, of frequencies which are integral multiples of $c/2l$, all propagated with a velocity c.

We shall consider two typical possibilities.

(i) *We might be given the values of* p *and* $\partial p/\partial t$ *as functions of* x *at a time* $t = 0$—that is to say we are to determine the subsequent motion of the air, given conditions at $t = 0$. At $t = 0$ we have from (24)

$$p_{t=0} = \sum_r B_r \sin\left(\frac{r\pi x}{l}\right), \quad \left(\frac{\partial p}{\partial t}\right)_{t=0} = \sum_r \frac{r\pi c}{l}\,C_r \sin\left(\frac{r\pi x}{l}\right) \;.\;. \quad (25)$$

so that, to determine B_r and C_r we have simply to *expand* $p_{t=0}$ and $(\partial p/\partial t)_{t=0}$ *as half-range sine series in* x by the process outlined in Art. 2 above. Substitution of these values in equation (24) then gives the complete answer to the problem.

(ii) *We might be given the value of* p *as a function of time at some definite position in the pipe,* $x = a$ (say). This case might arise if we imagined an organ pipe being blown in a particular way. We notice from (24) that, at any point in the pipe, p is a periodic function of the time, the period being $2l/c$, which is simply the time sound takes to travel twice the length of the pipe. (The physical meaning of this is that we may apply any pressures we like during this time $2l/c$, but that the subsequent history is controlled by reflections of the sound-waves at the ends of the pipe.) If we were given the variation of p at $x = x_0$ over one period, we could expand it as a Fourier series in t with this period, then the coefficients B_r and C_r can be obtained by putting $x = x_0$ in equation (24) and comparing coefficients.

If we were considering the problem of a pipe with a *closed* end at $x = 0$ and an open end at $x = l$, the boundary conditions would be $\partial p/\partial x = 0$ at $x = 0$ and $p = 0$ at $x = l$. Our method could be applied here too, but we should have to use in our x-expansion a set of functions that satisfied these boundary conditions individually. Such a set is

$$\cos\left(\frac{(2r+1)\,\pi x}{2l}\right) \text{— } r \text{ being an integer or zero.}$$

A series consisting of an expansion in terms of such functions we might almost call a "quarter-range" Fourier series, because x has to be increased by $4l$ before one of these terms repeats itself. It can be shown that these functions form an orthogonal set for integration over the range $x = 0$ to $x = l$, so that the calculation of the coefficients can proceed in a way quite analogous to that indicated in equations (14) or (15).

4. Application of Fourier analysis to the diffusion equation

On pp. 128, 228 we saw that diffusion in gases and liquids is governed by the law: rate of flow ∝ concentration gradient, which on p. 228 is shown to imply the equation

$$\lambda \frac{\partial^2 c}{\partial x^2} = \frac{\partial c}{\partial t} \quad \dots\dots\dots\dots\dots\dots \quad (26)$$

where λ is a constant.

A similar equation governs the conduction of heat, the propagation of electromagnetic waves in metals and the propagation of vortices in a liquid or gas flowing past a solid boundary (p. 152). Quite generally, an equation of this kind seems to govern *dissipative* processes, whereas the wave-equation governs *reversible* processes of energy conversion. Equation (25) can be solved by Fourier analysis in the same way as the wave-equation. The most usual problem that arises with diffusion is this. We are given a certain initial distribution of concentration, and wish to determine its subsequent behaviour as a function of x and t. Again, we may suppose that we are concerned with a vessel of length l, but that there is no physical reason why the concentration c should satisfy any particular boundary conditions at $x = 0$ and $x = l$. This would occur, *e.g.*, if the boundary were *porous* at $x = 0$ and $x = l$. If it were *impermeable* we should have to have $\frac{\partial c}{\partial x} = 0$ at $x = 0$ and $x = l$ because the flow at these points would have to be zero, and we should then have to expand c as a *half-range cosine series* in (x/l), because each term of such a series satisfies this boundary condition. For the former problem, we use an *ordinary* Fourier series in x which represents *any* function defined only in a finite range of the variable, rather than one that satisfies definite boundary conditions at the end of the range. Having decided this, we are led, by reasoning quite similar to that used in discussing equation (20), to seek a solution of (25) in the form

$$c = \tfrac{1}{2}a_0 + \sum_r \left[a_r \cos\left(\frac{2\pi r x}{l}\right) + b_r \sin\left(\frac{2\pi r x}{l}\right) \right] \quad \ldots\ldots \quad (27)$$

where the variation of c with time is allowed for by making the a's and b's functions of t, but not of x. Inserting (26) in (25), and applying the same result that there is no Fourier expansion of a null function, we get the following set of differential equations for the a_r, and a corresponding set for the b_r:—

$$\frac{da_r}{dt} + \frac{4\pi^2 r^2 \lambda}{l^2} a_r = 0 \quad \text{and} \quad \frac{da_0}{dt} = 0 \quad \ldots\ldots\ldots\ldots \quad (28)$$

The solution of these equations is

$$a_r = A_r e^{-\frac{4\pi^2 r^2 \lambda t}{l^2}}, \qquad b_r = B_r e^{-\frac{4\pi^2 r^2 \lambda t}{l^2}}, \qquad a_0 = \text{Constant}$$

where A_r and B_r are constants, so that the full solution of (25) is

$$c = \tfrac{1}{2}a_0 + \sum_r e^{-\frac{4\pi^2 r^2 \lambda t}{l^2}} \left[A_r \cos\left(\frac{2r\pi x}{l}\right) + B_r \sin\left(\frac{2r\pi x}{l}\right) \right] \ldots \quad (29)$$

If we are given the distribution of c as a function of x at $t = 0$, (29) shows us that a_0, A_r, B_r are just the Fourier coefficients of this distribution. The form of the exponential decay factors shows us that the rate of decay increases rapidly with r, which means that a distribution of concentration that varies rapidly with changing x equalises itself more quickly than one that only varies slowly with x. As t becomes large, the only term that survives is the $\tfrac{1}{2}a_0$, which means that the concentration has equalised itself. Equation (29) is of most use for relatively large values of t, and it may become almost useless for small t because a great many terms then have to be considered. Various other methods of solving the diffusion equation have been devised to deal with the case where t is small, but we shall not discuss them here. It is also possible to get solutions that are periodic in the time. These are of considerable interest in the *thermal* and *electromagnetic* applications of equation (26), but are scarcely relevant to the diffusion problem.

The process that we have used in solving equations (20) and (26), in which a set of orthogonal functions is used to enable us to reduce the solution of a partial differential equation to the problem of solving a set of *ordinary* differential equations is called the method of *separation of variables*, and it is probably the most important method of dealing with the partial differential equations of mathematical physics. It is applicable only to linear equations, but very many of the problems of mathematical physics are either inherently linear, as in much of electromagnetic theory, or can be made linear to a first approximation, as in acoustics.

5. The evaluation of certain integrals which occur in the kinetic theory of gases

In any discussion of transport properties of gases, it is necessary to average certain quantities over all possible values of molecular velocities, although it is sometimes possible to get correct results by making the incorrect assumption that all the molecules are moving at the same mean speed. The integrals usually wanted are of the type

$$I_n = \int_0^\infty c^n\, e^{-qc^2}\, dc,$$

where n is an integer, usually not very large, and q is a constant, usually $m/2kT$ in kinetic theory applications. We notice that I_n is elementary for $n = 1$. Put $c^2 = x$, $dx = 2c\,dc$, so that

$$I_1 = \int_0^\infty c e^{-qc^2}\, dc = \tfrac{1}{2}\int_0^\infty e^{-qx}\, dx = \frac{1}{2q}.$$

Differentiate both sides of this equation with respect to q

$$-\frac{\partial I_1}{\partial q} = \int_0^\infty c\,.\,c^2 e^{-qc^2}\, dc = I_3 = -\frac{\partial}{\partial q}\left(\frac{1}{2q}\right) = \frac{1}{2q^2}$$

$$\frac{\partial^2 I_1}{\partial q^2} = \int_0^\infty c^5 e^{-qc^2}\, dc = I_5 = \frac{\partial^2}{\partial q^2}\left(\frac{1}{2q}\right) = \frac{1}{q^3}$$

$$-\frac{\partial^3 I_1}{\partial q^3} = \int_0^\infty c^7 e^{-qc^2}\, dc = I_7 = \frac{3}{q^4}, \text{ etc.}$$

This process of differentiating the integrand to obtain a new integral is known as "differentiation under the integral sign". It is valid under conditions very similar to those that permit us to differentiate an infinite series term by term.

To evaluate I_n when n is even, we proceed as follows: We have

$$I_0 = \int_0^\infty e^{-qc^2}\,dc = \int_0^\infty e^{-qg^2}\,dg,$$

where g is another independent variable.

Therefore
$$I_0^2 = \int_0^\infty \int_0^\infty e^{-qc^2}\,dc\,.\,e^{-qg^2}\,dg,$$

the conditions for multiplying integrals being almost identical with those for multiplying two series term by term. (It is permissible here because both integrals are absolutely convergent.) To evaluate I_0^2, we put $c = r\cos\theta$, $g = r\sin\theta$, and the domain in which c and g may separately take any value between 0 and ∞ is simply that in which r may run from 0 to ∞ and θ from 0 to $\pi/2$. The element of area is $rdrd\theta$, so that (Fig. 109)

$$I_0^2 = \int_0^\infty \int_0^{\pi/2} rdrd\theta e^{-qr^2}$$

$$= \frac{\pi}{2}\int_0^\infty rdre^{-qr^2} = \frac{\pi}{2} I_1 = \frac{\pi}{4q}.$$

So that
$$I_0 = \tfrac{1}{2}\sqrt{\frac{\pi}{q}}.$$

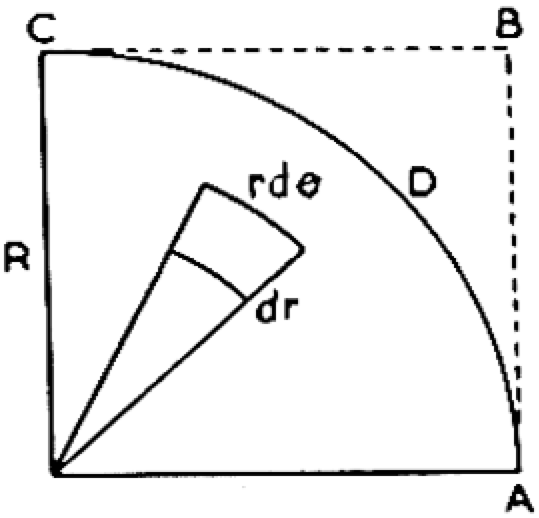


FIG. 109.

To justify this change of variables we have to show that it makes no difference to the integral. Now, if c and g run from 0 to R, the domain of integration is a square, whereas if we let r run from 0 to R and θ from 0 to $\pi/2$, the domain is a quadrant of a circle (Fig. 109). We shall set an upper limit to the error arising through omitting the crescent-shaped area $ABCD$, and show that this error tends to zero for large R.

For r sufficiently large ($> r_0$, say), the function re^{-qr^2} is a decreasing function of r, so that if R is taken greater than this previously assigned value r_0, the integrand is *larger* on the arc CDA than it is anywhere else in the area $ABCD$. Thus, we shall certainly overestimate the error if we replace the area $ABCD$ by the value $\frac{1}{2}R^2$ and the integrand by its maximum value in this area which is Re^{-qR^2} and multiply these two together. The error $< \frac{1}{2}R^3e^{-qR^2}$ which $\rightarrow 0$ as $R \rightarrow \infty$. Thus, our evaluation of I_0 is complete, and by differentiating it with respect to q we may obtain in succession I_2, I_4, I_6, etc.,

e.g.
$$I_2 = \tfrac{1}{4}\sqrt{\frac{\pi}{q^3}},\quad I_4 = \tfrac{3}{8}\sqrt{\frac{\pi}{q^5}},\quad I_6 = \tfrac{15}{16}\sqrt{\frac{\pi}{q^7}},\ \text{etc.}$$

As an example of the application of these integrals we shall compare the root mean square and the true mean speeds of molecules subject to a Maxwellian distribution.

For, according to Maxwell's assumptions, the number of molecules having a total velocity between c and $c + dc$ is proportional (p. 117) to

$$c^2 e^{-qc^2}\, dc, \quad \text{where } q = m/2kT.$$

Thus, the average total velocity in such an assembly is

$$\bar{c} = \frac{\displaystyle\int_0^\infty c^3 e^{-qc^2}\, dc}{\displaystyle\int_0^\infty c^2 e^{-qc^2}\, dc} = \frac{I_3}{I_2} = \frac{2}{\sqrt{\pi} q^{\frac{1}{2}}}.$$

On the other hand, the average of the square of the velocity is

$$\bar{c}^2 = \frac{\displaystyle\int_0^\infty c^4 e^{-qc^2} dc}{\displaystyle\int_0^\infty c^2 e^{-qc^2} dc} = \frac{I_4}{I_2} = \frac{3}{2q} = \frac{3\pi}{8}\,(\bar{c})^2.$$

Thus the root mean square and the mean speeds differ by a factor $(3\pi/8)^{\frac{1}{2}}$ for such a distribution, a difference of nearly 10 per cent.

BIBLIOGRAPHY

Carslaw, *Fourier Series and Integrals.* Macmillan, London (1930).

Margenau and Murphy, *The Mathematics of Physics and Chemistry.* Van Nostrand, New York (1943).

Piaggio, *Differential Equations.* Bell, London (1926).

Jaeger, *Introduction to Applied Mathematics.* O.U.P. (1951).

EXAMINATION QUESTIONS

N.S.T. = Cambridge Natural Sciences Tripos.
N.S.P. = Cambridge Preliminary Examination in Natural Sciences.
M.S.T. = Cambridge Mechanical Sciences Tripos.
B.Sc. (*Gen.*) = London General Examination for B.Sc.
B.Sc. (*Spec.*) = London Special Examination for B.Sc.

1. Explain what is meant by the "dimensions" of a physical quantity and find the dimensions of (*a*) the coefficient of viscosity, and (*b*) the gravitational constant (G).

On the assumption that the flow of liquid in a tube becomes turbulent at some critical velocity V which depends only on the viscosity η, the density of the liquid ρ and the radius of the tube r, find how V varies with these other quantities.

N.S.T., I.

2. Discuss critically our reasons for believing that the gravitational attraction between two point masses obeys the inverse-square law.

Show that the gravitational force due to the earth may be calculated as though all its mass were concentrated at its centre. Emphasise the assumptions that you make. *N.S.T., I.*

3. A long rectangular slab of glass of depth a, width b, and Young's modulus Y is bent in the plane containing a by the application of equal and opposite couples G applied at its ends. Show that it is bent into an arc of a circle and calculate its radius.

A flat piece of glass is placed on the upper surface of the bent piece and interference fringes are observed with light of wavelength λ. Show how Young's modulus can be deduced by using observations on these fringes to determine the radius of curvature of the bent slab in the main plane of bending. *N.S.T., I.*

4. Derive from kinetic theory an expression for the thermal conductivity of a gas. If the thermal conductivity of helium at N.T.P. is 3×10^{-4} cal. cm.$^{-1}$, sec.$^{-1}$, deg.$^{-1}$, what is (*a*) the velocity, and (*b*) the mean free path of a helium atom? The specific heat at constant volume is 0·75 cal. g.$^{-1}$ deg.$^{-1}$. The density of mercury is 13·6 g. cm.$^{-3}$, the density of helium at N.T.P. is $1{\cdot}8 \times 10^{-4}$ g. cm.$^{-3}$.

N.S.T., I.

5. Explain the principle of equipartition of energy showing the conditions under which the principle applies.

Describe briefly the most important cases where the predictions from the principle do not agree with experiment. In one of these cases show how the discrepancy has been accounted for by modern theory. *N.S.T., II.*

6. A gas is enclosed in a cylinder with a piston which moves out slowly with uniform velocity. On the assumption that the velocities of the gas molecules in three perpendicular directions are at each instant equal, and that the molecules are spheres which recoil from the moving piston elastically, show that the total

change dE in the kinetic energy E of the molecules is related to the change dV of the volume V by the relation

$$\frac{3}{2}\frac{dE}{E} = -\frac{dV}{V}$$

and show that this is consistent with the usual expression $pV^\gamma =$ constant, γ having the value appropriate to the simple gas model here assumed.

Comment on what would happen (*a*) if the molecules were polyatomic and (*b*) if the velocity of the piston were greater than the molecular velocities.

N.S.T., II.

7. Give an account of recent work on friction. *N.S.T., II.*

8. A beam of length l is encastered at one end and is in contact with a rigid support at the other end.

Show that the greatest sagging bending moment set up in the beam, as a load W travels slowly along it, occurs when the load is at a distance $\frac{1}{2}l(3 - \sqrt{3})$ from the encastered end. *M.S.T., I.*

9. A strut has a length $2l$ and is pin-jointed at its ends. For one-half of its length it has a cross-section with a moment of inertia I, the other half is of large section and can be treated as rigid. Show that the buckling load P for the strut is given by $al + \tan al = 0$, where $a^2 = P/EI$. *M.S.T., I.*

10. A flexible belt of total mass M and length $(\pi + 2)d$ passes round two solid cylinders, each of the same mass M and diameter d. The cylinders are then placed on a plane inclined at an angle of 30° to the horizontal. Find their velocity 1 sec. after release.

It can be assumed that the belt does not slip over the cylinders or over the plane, and that the friction between the cylinders at their point of contact is negligible. *M.S.T., I.*

11. The resultant force R resisting the motion of a projectile in earth is given by $R = A(a + bv^2)$ when A is the cross-sectional area of the projectile, v its velocity, and a and b are constants.

Find expressions giving

(i) the distance penetrated in terms of the mass and area of the projectile, its initial and final velocities, and the constants a and b;

(ii) the time of penetration in terms of the same quantities. For a certain material, a has the value $3{\cdot}33 \times 10^6$ poundals/sq. ft and b has the value 30 poundal-sec.^{2}/ft^4. Determine the delay necessary in the action of a fuse to produce detonation of a 500-lb. bomb of diameter 12 in. after a penetration of 15 ft, if it strikes with a velocity of 1,000 ft/sec.

M.S.T., I.

12. A homogeneous sphere of radius a is projected along a rough horizontal plane with velocity u, and angular velocity $6u/a$ in direction opposite to that required for rolling. Find an expression for the time taken for the sphere to return to its starting point, and the velocity with which it does so, if the coefficient of friction is μ. *M.S.T., I.*

13. Enunciate Raoult's law and show how it explains (*a*) the elevation in the boiling point of a dilute solution, and (*b*) the composition of the vapour phase in equilibrium with a binary liquid mixture which obeys the law. *N.S.T., I.*

14. Describe an accurate method of measuring the coefficient of viscosity of a gas, deriving from first principles the formula needed to determine the coefficient from the quantities observed. *N.S.T., I.*

15. Obtain an expression for the vapour pressure over a curved liquid surface, and explain in detail the phenomenon of "bumping" during boiling. *N.S.T., I.*

16. Relations between quantities obtained by means of a very simple picture of the processes involved are often of great assistance in the study of physical phenomena. Discuss this statement, illustrating your answer by reference to the behaviour of a gas at ordinary temperatures and pressures. (Assume that there are no forces between the molecules except at a collision, but that the molecules are of finite size.) *N.S.T., I.*

17. Explain clearly what is meant by surface tension and surface energy. By the use of thermodynamic arguments deduce a relation between these quantities, being careful to describe the physical meaning of each step in the theory. *N.S.T., I.*

18. Describe how the velocity of sound in a gas varies with temperature, pressure, and frequency. What information on the molecular structure of the gas may be obtained from measurements of these variations? *N.S.T., II.*

19. Discuss the assumptions that have to be made about the properties of materials in constructing a theory of elasticity and develop the argument to show that two elastic constants serve to define the behaviour of an isotropic solid.

Give a critical description of the theories developed to account for the properties of materials whose stress-strain curve is non-linear. *N.S.T., II.*

20. Compare critically three methods of measuring the gravitational constant (G).

Assuming that the moon describes a circular orbit of radius $3 \cdot 8 \times 10^5$ km. about the earth in 27 days, and that Titan describes a circular orbit of radius $1 \cdot 23 \times 10^6$ km. about Saturn in 16 days, compare the masses of the earth and Saturn. *N.S.T., I.*

21. Describe an accurate method for measuring the bulk modulus of elasticity of a liquid.

Prove from first principles that the value of Poisson's ratio for an isotropic solid cannot exceed $\frac{1}{2}$. *N.S.T., I.*

22. To what extent is it possible to correlate the mechanical properties of a substance with its atomic structure? *N.S.T., I.*

23. Discuss qualitatively the flow of gases through tubes as a function of pressure.

A vessel containing hydrogen at a temperature of 15° C. is connected by a tube of length 5 m. and radius 10 cm. to a fast pump. On the assumption that the pressure in the whole system is so low that intermolecular collisions may be neglected, calculate the velocity with which the gas streams into the tube. Find also the equilibrium pressure in the vessel if hydrogen leaks into it at a rate of 2 cm.^{3} (measured at 15° C. and 760 mm.) per minute.

$R = 8 \cdot 3 \times 10^7$ ergs deg.$^{-1}$ (g.-mol.)$^{-1}$. *N.S.T., II.*

24. Explain the connection between g, the acceleration due to gravity at the earth's surface, and G the gravitational constant.

Calculate the ratio of the mean densities of the earth and the sun from the following approximate figures:

Angular diameter of sun seen from earth = $\frac{1}{2}°$.
Length of 1° of latitude on earth's surface = 100 km.
One terrestrial year = $10^{7\cdot5}$ sec.
$g = 1{,}000$ cm. sec.$^{-2}$. *N.S.P., I.*

25. To what extent do you consider Newton's second and third laws of motion to be fundamental laws? Show how they lead to the principle of conservation of momentum in a collision between two bodies.

A nucleus, of mass 20 atomic mass units, is moving with a velocity of 5×10^8 cm. sec.$^{-1}$ when it breaks into two parts with an energy release of 10^{-4} erg. The heavier fragment has mass 16 units and is emitted at 90° to the original line of flight. Find the velocity of the lighter fragment.

(1 atomic mass unit = $1\cdot6 \times 10^{-24}$ g.) *N.S.P., I.*

26. Define the terms *surface tension* and *surface energy* and explain, without detailed analysis, why a soap-film gets cooler when it is stretched. Give the theory and experimental details of a method for studying the variation of surface tension of a liquid with temperature, and indicate graphically or analytically the kind of result you would expect. *N.S.P., I.*

27. Derive Bernoulli's equation,

$$p/\rho + gz + \tfrac{1}{2}v^2 = \text{constant}$$

for a non-viscous, incompressible fluid in streamline motion. Show how one can base on this result the design of simple devices for measuring (*a*) flow velocity, (*b*) volume rate of flow of water in a pipe. *N.S.P., I.*

28. Show that the gravitational attraction of a thin uniform spherical shell of material on a point mass placed anywhere within the shell is zero.

Assume the earth to be a perfect sphere, and that the density of the material is a function only of the distance from the centre of the sphere. Show that, for the case where the density decreases linearly with distance from the centre, the attraction of the earth on a point mass in a vertical mine shaft is independent of the depth provided that the density at the surface of the sphere is $\frac{1}{3}$ the density at the centre and that the depth of the mine is small compared with the earth's radius. Rotation of the earth is to be neglected. *N.S.T., I.*

29. By considering the motion of a particle on the heavy rim of a wheel otherwise light, which is rotating rapidly about an axis through its centre and perpendicular to its plane, explain qualitatively why a couple applied perpendicular to that axis will cause the wheel to precess about an axis perpendicular to both the axis of rotation of the wheel and the axis of the couple. Hence, or otherwise, show that the couple C is given by the relation $C = I\omega\Omega$, where I is the moment of inertia of the wheel, ω is the angular velocity of rotation of the wheel, and Ω is the angular velocity of precession.

Describe one practical application of the gyroscope, emphasising any modifications of a simple rotating wheel gyroscope which make it suitable for the application you describe. *N.S.T., I.*

30. The inner end of a flat spiral spring is attached to an axle fitted to a wheel, the outer end being fixed to the framework to which the (frictionless) bearings of the axle are fixed. A small sinusoidal couple of frequency n is applied to the axle. Discuss the motion of the wheel at various frequencies. How would the motion be affected if a light brake which exerted a couple on the axle proportional to the angular velocity of the axle were applied?

If the wheel is a disc 10 cm. in diameter weighing 1,000 g., the couple exerted by the spring 10,000 dyne cm. per radian deflection, and the couple exerted by the brake 10,000 dyne cm. per radian per second, what is the maximum amplitude of oscillation of the wheel which can be produced by an applied sinusoidal couple of amplitude 5,000 dyne cm. and at what frequency does this occur? *N.S.T., I.*

31. What is meant by a perfectly elastic substance? In what ways do practical materials differ from a perfectly elastic substance?

You are given a steel rod of circular cross-section, 50 cm. long and 1 cm. in diameter. How would you determine one of its elastic constants? Estimate the magnitude of the quantities you would measure.

(Young's modulus $= 2 \times 10^{12}$ dyne cm.$^{-2}$. Rigidity modulus $= 1 \times 10^{12}$ dyne cm.$^{-2}$.)

N.S.T., I.

32. By considering *either* a stretched string *or* a gaseous or solid medium, show that the velocity of propagation of waves along the string or in a medium can be calculated from properties of the string or medium which can be measured by methods not involving wave motion.

How would you find experimentally the velocity of sound along a metal bar?

N.S.T., I.

33. How would you illustrate experimentally that an actual gas does not behave exactly as a perfect gas? Show how such behaviour is accounted for by the kinetic theory, illustrating your answer by deriving van der Waals' equation by an elementary method.

A gram-molecule of gas, for which the intermolecular attraction is known to be negligible, has a volume of 2,486 cm.^{3} at a pressure of 10 atmospheres. At a pressure of $\frac{1}{10}$ atmosphere, and the same temperature, the volume is 246,000 cm.^{3} and at this and lower pressures the gas can be treated as a perfect gas without appreciable error. Obtain an approximate value of the diameter of a molecule of gas on the assumption that Avogadro's number is 6×10^{23} molecules per gram-molecule.

N.S.T., I.

34. State what you understand by the *vapour pressure of a liquid*. Discuss the phenomena associated with vapour pressure from the standpoint of the kinetic theory of matter explaining (*a*) how under fixed conditions a constant pressure of vapour in contact with a liquid is maintained, (*b*) the effect of change of temperature, (*c*) the effect of introducing a perfect gas into the volume containing the liquid and vapour.

Liquid mercury at 0° C. is contained in a highly-evacuated vertical cylinder of cross-section 5 sq. cm. Immediately above, but not in contact with, the mercury surface, is a surface cooled by liquid air, so that mercury atoms leaving the liquid surface are condensed, and cannot return to the liquid mercury surface. It is found that 3 g. of liquid mercury evaporate in one hour. Calculate an *approximate* value of the vapour pressure of mercury at 0° C.

(Mean velocity of mercury vapour atoms at 0° C. $= 6{,}000$ cm. sec.$^{-1}$.)

N.S.T., I.

35. Give an account, without mathematical details, of the characteristic behaviour of resonant systems, drawing examples from as wide as possible a range of physical phenomena. *N.S.P., II.*

36. Show that Poisson's ratio for an elastic solid must lie between $\frac{1}{2}$ and -1.

For a certain material the velocity of longitudinal waves is greater in an unbounded medium than in a thin rod by the factor $\sqrt{6/5}$. Determine the possible values of Poisson's ratio. *N.S.P., II.*

37. Discuss the justification for speaking of the "size" of a neutral atom or molecule. Describe *three* ways, as widely different as possible, by which molecular diameters can be estimated. Quote or derive any relevant formulae. *N.S.P., II.*

38. Show to what extent the crystal structure may determine the strength and other mechanical properties of solids.

Review other factors which must be taken into account in the theoretical treatment of the problem of the mechanical properties of solids, and discuss one of these factors in detail. *N.S.T., II.*

39. In Heyl's method of measuring the constant of gravitation G, two rigidly-linked spheres are suspended by a torsion fibre, and oscillate in the gravitational field of two fixed spheres. Discuss from first principles how you would realise this method in practice, paying particular attention to the choice of dimensions and materials of all parts of the apparatus, and to the difficulties you would expect to encounter. *N.S.T., II.*

40. Discuss the importance of the concept of mean free path in the kinetic theory of gases.

Give an account of methods of producing and measuring very low pressures, paying particular attention to the practical limitations of each method you describe. *N.S.T., II.*

41. Give a general account of what may be learnt from measurements at sea of g, the acceleration due to gravity, and indicate the accuracy necessary to give useful information.

Explain why, even in a perfectly calm sea, the apparent value of g measured in a moving submarine differs from the true value and indicate how the difference may be estimated. *N.S.T., II.*

42. What methods are available for calculating the way in which sound waves of finite amplitude are propagated in a medium in which the velocity of sound varies as the nth power of the density. Such a medium is initially at rest in a rigid pipe; outline a practicable computation procedure for determining the propagation of disturbances produced in the medium as a result of a prescribed motion of a piston at one end. Derive any equations you need.

At what stage, if any, would you expect your process to break down? What modifications of procedure would enable you to follow the history of the disturbance beyond that stage? *N.S.T., II.*

43. Discuss the analysis of a non-periodic function into sinusoidal components and give an account of some applications of such an analysis to physical problems. *N.S.T., II.*

44. Derive an expression for the torsional rigidity of a hollow circular cylinder of length l, internal radius a, and external radius b.

A propeller shaft has to possess a certain torsional rigidity. Find what percentage of material can be saved by making the shaft hollow instead of solid, if the internal radius of the hollow shaft is half the external radius. *N.S.T., II.*

45. Give some account of the methods used in making a gravitational survey. *N.S.T., II.*

46. Show from first principles that the couple G acting on a rigid body free to rotate about a fixed axis is given by $G = d/dt\,(I\dot{\theta})$ where I is the moment of inertia of the body about the axis and $\dot{\theta}$ is the angular velocity of the body.

A reel of thread of moment of inertia Mk^2 about its axis is allowed to unwind under gravity. If the radius of the reel is a, find the acceleration of the reel and the tension in the threads. Indicate, without solution, the form which the equations of motion would take if the mass of the thread were not negligible. *N.S.T., I.*

47. Give an account of the main features of sliding friction, and the effect of lubricants in reducing friction between sliding surfaces.

A rope is wound round a fixed cylinder so as to make n complete turns. Show that if one end of the rope is held by a force T, a force $Te^{2\pi n\mu}$ must be applied to the other end to produce slipping where μ is the coefficient of friction between rope and cylinder.

Find also the work necessary to turn the cylinder through one complete turn under these conditions. *N.S.T., I.*

48. How may molecular diameters be estimated by methods based on the kinetic theory of gases?

A parallel beam of electrons passes through a monatomic gas at a pressure of 10^{-4} mm. of mercury. Assuming that an electron is removed from the beam whenever it approaches within 2×10^{-8} cm. of the centre of an atom of the gas, find the reduction in intensity of the electron beam in passing through a metre of gas.

[Avogadro's number $= 6{\cdot}06 \times 10^{23}$ (g.-mol.)$^{-1}$.] *N.S.T., I.*

49. What is the modulus of rigidity? Describe how it may be measured for a material available as a fine wire.

A body whose moment of inertia is I hangs from a wire for which a couple C is required to turn the ends relative to one another by 1 radian. The upper end of the wire is suddenly twisted through an angle θ. When the body has reached the limit of its swing, the wire is suddenly twisted back to its original position. Derive, as shortly as possible, the angular velocity of the body as it passes through its original position. *N.S.P., I.*

50. A long needle of radius r and density ρ floats on the surface of a liquid of density ρ_1. The surface tension of the liquid is S, and the angle of contact with the needle is θ. Derive expressions from which the distance of the needle's axis from the liquid surface could be found. (An explicit expression for this distance need not be given.) Is there any sense in which the weight of the needle is equal to that of the liquid displaced? *N.S.P., I.*

51. Prove that the motion of the centre of gravity of a rigid body under external forces can be determined by supposing that the mass of the body is concentrated at the centre of gravity and that the forces are acting there.

By considering the case of rotation about the axis OZ, with all the external forces perpendicular to this axis, establish the principle of conservation of angular momentum.

A piece of uniform metal wire is bent into a circle of radius a and is supported so that it can rotate about a horizontal axis at right angles to its plane, passing through the centre. An arc of length l is cut from the circle. Find the period of small oscillations about the position of equilibrium. *B.Sc.* (*Gen.*).

52. Explain how to find Young's modulus for copper provided in the form of (*a*) a wire about 20 cm. long and about 1 mm. radius, and (*b*) a rod of about 1 m. long and of rectangular cross-section about 2 cm.$\times$0·5 cm. Give experimental details in both cases and an account of the theory in one of them.
B.Sc. (*Gen.*).

53. A capillary tube of length l and radius a is fixed vertically through the base of an upright vessel. Water in the vessel is maintained at depth h while it flows through the tube. Find the rate of flow through the tube in cubic centimetres per second, assuming that the water leaves the tube with small velocity, the length of the tube in the water being l'.

Explain how this result can be applied in viscosity measurements.
B.Sc. (*Gen.*).

54. Derive the expression for the bending moment of a beam bent in the form of an arc of radius R.

A light rod is held horizontally at one end and carries a load at the other. It is of circular section but varying radius. Find an expression for the stress at the surface of the rod in terms of the load, radius r of the section, and distance d from the loaded end.

Show that this surface stress is constant all along the rod provided that r is proportional to $d^{\frac{1}{3}}$, and discuss a possible application of this result. *B.Sc.* (*Gen.*).

55. Describe how the coefficient of viscosity of a liquid can be determined by the method of flow through a narrow tube. Derive the formula required for the determination. *B.Sc.* (*Gen.*).

56. Give an account of the work of Andrews on the investigation of the pressure-volume relation of fluids, *e.g.* carbon dioxide. Discuss the agreement of the experimental results with the equation of state proposed by van der Waals.
B.Sc. (*Gen.*).

57. Describe a convenient method for the investigation of the variation of surface tension with temperature.

Two plane rectangular glass plates are hinged at one edge. They are placed with this edge vertical and the lower horizontal edges in water to a small depth. If the sides of the plate are long and their inclination very small, determine the shape of the curve made against each plate by the water surface between them.
B.Sc. (*Gen.*).

58. Define the four ordinary elastic constants of a homogeneous solid.

If Young's modulus for a specimen of copper is $1{\cdot}23\times10^{12}$ dynes/cm.^{2} and the modulus of rigidity is $4{\cdot}55\times10^{11}$ dynes/cm.^{2} find Poisson's ratio for the metal, deducing the necessary formulae. *B.Sc.* (*Gen.*).

59. How would you devise and carry out an experiment to ascertain whether the presence of air in contact with the surface affects the surface tension of water?

In general, does the nature of the fluid in contact with the surface make any difference to the surface tension? *B.Sc. (Gen.).*

60. A solid sphere of radius r is hung as a pendulum by a light flexible wire, the centre of the sphere being distant l from the point of suspension. Find, from first principles, an expression for the error in the period made by assuming that the mass of the sphere can be considered as concentrated at its centre.

If the diameter of the sphere is $0{\cdot}1l$, what is the approximate percentage error? *B.Sc. (Gen.).*

61. A couple C, applied statically to the lower end of a wire whose upper end is fixed, twists it through ϕ degrees. Another wire of the same material, twice as long and twice as thick as the first wire, carries at its lower end a disc of metal, of density ρ, radius r, and thickness d, rigidly attached so that its axis coincides with that of the wire. Find an expression for the time period of the oscillations of the disc. Any formulae used should be proved. *B.Sc. (Gen.).*

62. Define *shearing strain.* Describe and explain, with illustrative diagrams two ways in which it may be produced.

A metal tube, 1 m. long and of internal and external radii 4·0 cm. and 4·2 cm. respectively, is clamped at one end. Find the magnitude of the couple which must be applied to the other end to twist it through 30°. Derive the expression used in the calculation. The modulus of rigidity of the metal is $2{\cdot}5 \times 10^{11}$ dynes/cm.$^{-2}$. *B.Sc. (Gen.).*

63. Discuss the motion of a particle in a central attractive field of force which varies according to the inverse-square law. Explain how the form of the path depends on the initial conditions and, applying your discussion to the motion of a planet, establish Kepler's laws. *B.Sc. (Spec.).*

64. Derive from first principles, an expression for the depression at the mid-point of a beam of uniform mass per unit length, loaded at the centre and supported at the ends on knife edges at the same level.

In the case of a beam of negligible mass, find the period of small oscillations when the load is slightly displaced vertically from the position of equilibrium and then released. *B.Sc. (Spec.).*

65. A large drop of liquid lies on a horizontal surface in air. Take the origin at the centre of the upper surface and show that at any point P in a vertical plane section the ordinate y and the slope ψ of the tangent to the intersection of the plane and surface are related by the equation

$$g\rho y^2 = 2T\,(1-\cos\psi)$$

where ρ is the density of the liquid and T its surface tension. Assume that the curvature in the horizontal section through P is very small.

Describe and explain how this formula can be applied to determine the surface tension and angle of contact in certain cases. Discuss briefly the measurement of angles of contact. *B.Sc. (Spec.).*

66. Give an account of the methods which have been used to measure *either* (*a*) osmotic pressures *or* (*b*) coefficients of diffusion. *B.Sc. (Spec.).*

67. Explain briefly the details of the process of producing a very high vacuum in a vessel of about 1 litre capacity starting at atmospheric pressure and explain the action of a diffusion pump as fully as you can. Illustrate both parts of your answer by diagrams. *B.Sc.* (*Spec.*).

68. Give an account of the measurement of very low pressures. *B.Sc.* (*Spec.*).

69. A particle which is in motion under the action of a force directed towards a centre and proportional to its distance from it is in a medium producing a resistive force proportional to the velocity and is, in addition, acted upon by a simple harmonic force.

Investigate the nature of the motion from the instant of application of the simple harmonic force. Consider the response of the particle to the force and show how the resistive force may be determined. Give examples of the application of this problem. *B.Sc.* (*Spec.*).

70. A particle of mass m, initially at rest, is acted on by a constant force fm and by a resisting force $mk\dot{x}$ where $\dot{x}$ is the velocity of the particle at a distance x from the origin. Derive expressions for (*a*) the terminal velocity acquired, (*b*) the velocity at a time t after application of the force, (*c*) the distance travelled by the particle in this time.

Examine the case when the constant force is replaced by a sinusoidal force $mA \sin pt$.

An oil drop, of density 0·95 g. cm.$^{-3}$ and radius 10^{-4} cm. and carrying a charge of e e.s.u., is driven vertically upwards through air of density 0·0013 g. cm.$^{-3}$ by an electric field of 2,000 V. cm.$^{-1}$ with a constant velocity of 0·036 cm. sec.$^{-3}$. Calculate the magnitude of e, being given that the viscosity of air is 181×10^{-6} g. cm.$^{-1}$ sec.$^{-1}$. *B.Sc.* (*Spec.*).

71. A light uniform cylindrical rod, 100 cm. long and 1 cm. diameter, is supported symmetrically on two knife edges 50 cm. apart in a horizontal plane. Equal masses of 1 kg. are suspended from the centre and ends of the rod. Calculate (*a*) the depressions at the centre point and at the ends of the rod, (*b*) the strain energy stored in the rod. Young's modulus of the material of the rod is 10×10^{10} dynes cm.$^{-2}$. *B.Sc.* (*Spec.*).

72. Give the theory and practical details of a method of determining Poisson's ratio for glass by measurements of interference fringes. *B.Sc.* (*Spec.*).

73. Liquid flows in stream-line motion in a horizontal tube of length l and internal radius a. Derive an expression for the volume passing in unit time due to an effective pressure-difference p between the ends of the tube. Show that, if a cylindrical coaxial rod of radius $a/2$ and length l is inserted into the tube, the rate of flow is reduced by about 88 per cent. for the same pressure-difference between the ends. *B.Sc.* (*Spec.*).

74. Give a brief account of the principal methods of determining Avogadro's number and describe in detail one method based on Brownian motion in liquids. *B.Sc.* (*Spec.*).

75. Give an account of the techniques employed in the production and measurement of very high pressures. *B.Sc.* (*Spec.*).

76. Assuming that the molecules of a gas are all moving with a random velocity c and that there are N molecules per c.c. derive an expression for the mean free path of a molecule.

Deduce an expression for the viscosity of a gas in terms of the mean free path of a molecule and discuss the dependence of the viscosity on pressure and temperature.

B.Sc. (*Spec.*).

77. Derive an expression for the difference in the saturation pressure of water vapour over a spherical surface and a plane surface.

Discuss the theory of drop formation in a Wilson cloud chamber.

B.Sc. (*Spec.*).

78. Write down the equation of motion for a particle under a restoring force proportional to the displacement from a fixed point, a resistance equal to $k \times$ velocity and an external simple harmonic force. Solve the equation and, in the steady state, if W_0 and f_0 denote the kinetic energy and frequency at resonance, W and f these quantities in any condition, show that

$$\sqrt{\frac{W_0 - W}{W}} = \frac{2\pi\,(f_0^2 - f^2)}{kf}.$$

Give an example of the application of this theory. *B.Sc.* (*Spec.*).

79. Derive the equations of motion for a perfect fluid under body forces.

When a velocity potential exists and the motion is steady show that:

$$\int \frac{dp}{\rho} + V + \tfrac{1}{2}v^2 = \text{constant}$$

where p denotes the pressure, ρ the density, V the potential energy per unit mass, and v the velocity of the fluid.

Apply this result to the case of a perfect gas which flows adiabatically through an orifice in the wall of a containing vessel, the internal and external pressures being P and B respectively and deduce the velocity of efflux. *B.Sc.* (*Spec.*).

80. Give an account of the methods used and the principles involved in the determination of the bulk modulus of liquids. *B.Sc.* (*Spec.*).

81. A piece of glass tubing of about 4 mm. diameter is joined at one end to a length of capillary tubing and placed vertically with the capillary tubing below. A pellet of mercury in the wider tube falls slowly with velocity v, driving air through the capillary.

The length of the mercury column is such that it would produce a pressure p, the atmospheric pressure is B and to allow for the friction between the pellet and the tube the effect of the pressure due to the pellet must be diminished by an amount c. Show that

$$\left(p - \frac{p^2}{2B} - c\right)\Big/ v = \text{constant}$$

assuming c/B to be negligible.

Determine the value of the constant in terms of the appropriate quantities and explain how the viscosity of air could be determined by the method suggested by this question.

B.Sc. (*Spec.*).

82. Derive the expression $v = \sqrt{\dfrac{K + 4n/3}{\rho}}$ for the velocity of elastic waves in a medium in which there are no lateral changes of dimension.

Apply the result to the case of sound-waves in air. *B.Sc.* (*Spec.*).

TABLES

The following tables are only intended to be a rough guide to the student in working out problems. For more accurate values reference should be made to published tables. Some of the quantities listed, such as the elastic constants, vary considerably from specimen to specimen.

TABLE I

FUNDAMENTAL CONSTANTS

Planck's constant, h	$6{\cdot}6 \times 10^{-27}$ erg sec.
Charge on the electron, e	$4{\cdot}8 \times 10^{-10}$ electrostatic units
Mass of the electron, m	$9{\cdot}0 \times 10^{-28}$ g.
Mass of the hydrogen atom, M	$2{\cdot}0 \times 10^{-24}$ g.
Velocity of light, c	$3{\cdot}0 \times 10^{10}$ cm. sec.
Gravitational constant, G	$6{\cdot}7 \times 10^{-8}$ C.G.S. units
Avogadro's number, N	$6{\cdot}0 \times 10^{23}$ mol./g. mol.
Radius of first orbit of hydrogen atom, $\frac{h^2}{4\pi^2 me^2}$	$4{\cdot}0 \times 10^{-9}$ cm.
Boltzmann's constant, k	$1{\cdot}4 \times 10^{-16}$ erg/° K.

TABLE II

PROPERTIES OF THE EARTH

Mass	$6{\cdot}0 \times 10^{27}$ g.
Mean density	$5{\cdot}5$ g./cm.^{3}
Mean density of crust	$2{\cdot}6$ g./cm.^{3}
Acceleration of gravity (England)	980 cm. sec.$^{-2}$
Volume of earth	10^{27} cm.^{3}
Volume of ocean	$1{\cdot}5 \times 10^{24}$ cm.^{3}
Area of land	$1{\cdot}5 \times 10^{18}$ cm.^{2}
Area of ocean	$3{\cdot}7 \times 10^{18}$ cm.^{2}

TABLE III

DENSITIES AND ELASTIC CONSTANTS OF SOME COMMON SUBSTANCES

	DENSITY g./c.c.	BULK MODULUS dynes/cm.^{2}	YOUNG'S MODULUS dynes/cm.^{2}	RIGIDITY MODULUS dynes/cm.^{2}	POISSON'S RATIO dynes/cm.^{2}
Aluminium	2·7	$7{\cdot}5 \times 10^{11}$	$7{\cdot}1 \times 10^{11}$	$2{\cdot}6 \times 10^{11}$	0·33
Copper	8·9	$1{\cdot}3 \times 10^{12}$	$1{\cdot}2 \times 10^{12}$	$4{\cdot}6 \times 10^{11}$	0·35
Steel	7·8	$1{\cdot}6 \times 10^{12}$	$2{\cdot}1 \times 10^{12}$	$8{\cdot}3 \times 10^{11}$	0·28
Cast Iron	7·8	$9{\cdot}6 \times 10^{11}$	$2{\cdot}1 \times 10^{12}$	$\sim 4{\cdot}0 \times 10^{11}$	0·25
Lead	11·4	$5{\cdot}0 \times 10^{11}$	$1{\cdot}6 \times 10^{11}$	$0{\cdot}6 \times 10^{11}$	0·45
Crown Glass	2·5	$5{\cdot}0 \times 10^{11}$	$7{\cdot}0 \times 10^{11}$	$3{\cdot}0 \times 10^{11}$	0·23
Brass	8·5	$1{\cdot}1 \times 10^{12}$	$1{\cdot}0 \times 10^{12}$	$3{\cdot}5 \times 10^{11}$	0·37
Water	1·0	$2{\cdot}0 \times 10^{10}$	—	—	—
Mercury	13·6	$2{\cdot}6 \times 10^{11}$	—	—	—

TABLE IV

BREAKING STRESSES

	Breaking Stress dynes/cm.^{2}
Aluminium	$0{\cdot}6 \times 10^9$
Copper (cast)	$1{\cdot}5 \times 10^9$
„ (rolled)	$2{\cdot}5 \times 10^9$
Steel (mild)	$4{\cdot}5 \times 10^9$
„ (high tensile) ...	$12{\cdot}0 \times 10^9$
Lead	$0{\cdot}16 \times 10^9$
Brass	$2{\cdot}0 \times 10^9$

TABLE V

PROPERTIES OF SOME LIQUIDS

Liquid	Density	Surface Tension	Viscosity (15° C.)
Water	1·0	72—80	0·011
Methyl alcohol	0·8	23	0·006
Ethyl alcohol	0·8	22	0·012
Ether	0·7	17	0·002
Benzene	0·9	29	0·007
Glycerine	1·26	63	15·000
Turpentine	0·8	27	0·015
Carbon tetrachloride ...	1·6	26	0·001
Mercury	13·6	465	0·015

TABLE VI

VARIOUS ESTIMATES OF MOLECULAR DIAMETERS (UNIT 10^{-8} cm.)

	From Quantum Mechanics	From Viscosity of Gas	From van der Waals' Constant b	From X-ray Scattering	From Density of Liquid (upper limit)
Hydrogen	3·0	2·7	2·5	—	3·9
Helium	2·6	2·2	2·0	—	3·6
Neon	2·8	2·4	2·7	1·3	3·8
Argon	3·2	3·7	2·9	2·1	4·0
Krypton	—	4·1	3·2	2·4	4·5
Xenon	—	4·9	3·5	2·7	4·5
Oxygen	—	3·7	2·9	—	3·7
Nitrogen	—	3·8	3·6	—	4·0

The column headed "From Quantum Mechanics" is obtained from theoretical estimates of the interactions between molecules. The "diameter" of a molecule is taken as that distance at which attractive and repulsive energies between two molecules are numerically equal. The column headed "From X-ray Scattering" is taken from W. L. Bragg's paper, *Phil. Mag.*, 1920, **40**, 169, and probably represents the diameters of the outer electronic orbits. There are theoretical reasons for believing that, for light molecules such as hydrogen and helium, the density of the liquid should be somewhat lower than the value to be expected if the molecules are in contact.

ANSWERS TO EXAMPLES

Chapter II (p. 44)

3. $M_2 + m_2 = 1{\cdot}01M_1 + 1{\cdot}01m_1 + 0{\cdot}01m_3$ **4.** $y + gx^2/2v_0^2 = v_0^2/2g$.

5. $M = \frac{1}{3}\pi\rho l^3 \tan^2\theta$, $I = \frac{1}{10}\pi\rho l^5 \tan^4\theta$. **6.** $(2g/3)^{\frac{1}{2}}$. **8.** $\theta = \tan^{-1} 1{\cdot}3$.

9. Ratio $= \left[\dfrac{a + (a^2 - b^2)^{\frac{1}{2}}}{a - (a^2 - b^2)^{\frac{1}{2}}}\right]^2$ or $\dfrac{16a^2}{b^2}$ if $b \ll a$.

10. (*a*) Critical velocity $v = \sqrt{gR}$, where R is radius of earth. Most favourable direction is tangential to earth's surface.

(*b*) Critical velocity $v = \sqrt{2gR}$. Direction is immaterial.

Chapter III (p. 71)

2. Mass of sun about 2×10^{33} g. Mass of earth about 6×10^{27} g.

3. 33 km. approx.

Chapter IV (p. 111)

5. Length/Radius $= 10\pi$. **6.** $\dfrac{dR}{R} = \dfrac{dl}{l}(1 - 2\sigma)$.

7. (*a*) 5×10^5 cm./sec. (*b*) $5{\cdot}9 \times 10^5$ cm./sec. **9.** About 3 parts in 10^8.

Chapter V (p. 137)

1. 2×10^5 cm./sec.

2. Free path $\sim 1/N\pi\sigma^2$. Distance between neighbours $\sim 1/N^{\frac{1}{3}}$ where $N =$ number of molecules per unit volume. Ratio approaches unity if molecules are nearly in contact.

Chapter VI (p. 181)

2. Virtual mass is equal to mass of displaced fluid.

6. See p. 169. **7.** See p. 160.

Chapter VIII (p. 230)

3. $5{\cdot}2 \times 10^{-2}$ cm./sec.

Chapter IX (p. 275)

1. For charged drop, total energy $=$ constant $+ S.4\pi r^2 - \dfrac{L}{\rho}.\dfrac{4\pi r^3}{3} - \dfrac{e^2}{2r}$. The equilibrium value of r is obtained by putting the derivative of this expression with respect to r equal to zero.

2. $5{\cdot}5 \times 10^{-5}$ °C.

Chapter X (p. 284)

(In the formulae given below, f and ϕ stand for arbitrary functions.)

1. $V = \sqrt{g\lambda}\,\phi\,(\lambda/h)\,f\,(T/\rho g\lambda^2)$. **2.** See p. 255. **3.** $V = \left(\dfrac{P}{\rho}\right)^{\frac{1}{2}} f\left(\dfrac{a}{\lambda}\right)$.

4. $G = Wl\phi\left(\dfrac{Ya^3}{Wl}\right) f\left(\dfrac{a}{l}\right)$. **5.** $h = \dfrac{Y}{\rho g} f\left(\dfrac{Y}{\rho g a}\right)$. **6.** $m \propto \dfrac{R^3}{GT^2}$.

HINTS AND ANSWERS TO EXAMINATION QUESTIONS

(p. 303)

1. $G \propto [M]^{-1}[L]^3[T]^{-2}$, $\eta \propto [M][L]^{-1}[T]^{-1}$. $\Gamma \propto \eta/\rho r$.

3. Radius $= Yab^3/12G$. See p. 90.

4. (*a*) $1{\cdot}3 \times 10^5$ cm./sec. (*b*) 5×10^{-5} cm.

6. **Hint.**—Use the principle of conservation of momentum to estimate the change in mean K.E. when the piston is moving at a definite small velocity v for a time t. Then let $v \to 0$ but keep vt finite.

8. Encastered means the same as clamped or "built-in".

10. **Hint.**—Use conservation of energy. Since system is uniformly accelerated from rest, linear velocity v of spheres down plane is proportional to t, distance run x to t^2. Angular velocity of spheres $= 2v/d$. Upper straight portion of belt moves with velocity $2v$. Determine K.E. of each small curved portion of belt separately and integrate. Energy equation now gives a relation between v and x.

11. 0·02 sec.

12. Time $= 2u/\mu g$, Velocity $= u$. **17.** See p. 232.

20. Ratio is $\dfrac{(3{\cdot}8 \times 10^5)^3}{(1{\cdot}23 \times 10^6)^3} \times \dfrac{16^2}{27^2} = 10^{-2}$ approx.

23. **Hint.**—The examiner obviously did not expect an elaborate derivation of the formula for rate of flow, complete with numerical coefficient. Something on the lines suggested on p. 131 would be appropriate, which would enable an "order of magnitude" calculation to be made.

24. **Hint.**—A very rough calculation is sufficient. The examiner is trying to test your physical insight. Actual value of ratio is 0·26, the sun having a density of 1·4 and the earth 5·5 g./c.c.

25. $5{\cdot}1 \times 10^9$ cm./sec. **30.** $\sqrt{5}/4$ radians at $4\pi/\sqrt{5}$ oscillations/sec.

33. **Hint.**—Use van der Waals' equation in the form

$$P(v - b) = RT.$$

Diameter $= 2{\cdot}2 \times 10^{-8}$ cm.

34. 2×10^{-4} mm. of mercury or 0·3 dynes/cm.².

36. See Hint on Example 2, p. 111. Possible values of σ, $\frac{1}{4}$, or $-\frac{1}{3}$.

39. **Hint.**—This is not meant to be merely a descriptive question. Begin by computing, in order of magnitude, the expected torque due to the presence of two fixed spheres of practicable size. Compare this with the torque to be expected in a torsion fibre of sufficient tensile strength and so estimate the fractional change in period that you expect the presence of the fixed spheres to produce. Consider how you will measure this change. (The examiner would not mind if your apparatus differed substantially from what was actually used, provided that your reasoning is sound.) Most moduli of elasticity of metals are of the order of $10^{11} - 10^{12}$ dynes/cm.^{2}. Elastic limits will be about $10^8 - 10^9$ dynes/cm.^{2}. $G = 6{\cdot}7 \times 10^{-8}$ C.G.S. units.

42. Hint.—See pp. 175-7. The Riemann equations can be solved numerically by step-by-step methods.

43. Hint.—Note the words "non-periodic". This means that we are interested in the limiting form of Fourier analysis when the fundamental period becomes very long. This is not actually dealt with in this book, but is considered in the more advanced treatises listed on p. 302.

44. Hint.—See p. 87. Saving = 22 per cent.

46. Acceleration $= ga^2/(k^2 + a^2)$, Tension $= Mgk^2/(k^2 + a^2)$.

48. About 40 per cent. reduction. (Nature of the theory does not justify giving a more accurate result.)

51. Hint.—Formula for period of compound pendulum is given on p. 31. In this we must write a for k, and the distance of the centre of gravity of the arc from the centre of the circle for l.

53. See p. 147.

54. Hint.—See p. 89. The surface stress can be obtained from the elongation of a fibre on the outer surface, which can be determined geometrically. See Fig. 36.

55. See p. 147. **58.** 0·35.

60. Period increased by factor $\left(1 + \frac{2}{5}\frac{a^2}{l^2}\right)^{\frac{1}{2}}$. Error $\frac{1}{5}$ per cent.

61. Period $= 2\pi\,(40\rho r^4 d/c)^{\frac{1}{2}}$. (Note that ϕ is in degrees.) **62.** See p. 87.

63. See p. 21. **64.** See p. 87. **65.** See p. 253. **69.** See p. 285.

70. See pp. 16 and 285. **71.** See p. 87. **72.** See p. 98.

73. Hint.—A slight modification of the work on p. 147 is called for, the velocity now having to vanish at *both* the inner rod and outer wall. In the text we put a constant of integration equal to zero when integrating equation [5(*a*)], p. 148, in order to avoid a singularity at $r = 0$, but this constant now enables us to satisfy the extra boundary condition.

76. See p. 123. **77.** See p. 243.

INDEX

ACCELERATION equation of hydrodynamics, 142
— of gravity by pendulum measurements, 50, 54
Accommodation coefficient, 134, 267
Acoustic approximation in hydrodynamics, 156
Adam, 257
— measurements of angle of contact, 258
Adams, 243
Adsorbed films, mobile and immobile, 268, 269
Adsorption, 267
— critical, 272
— isotherm, 269
Advance of perihelion, 27, 68, 69
Airy's mine experiment, 51
Altitude, effect of, on gravitational acceleration, 66
Angle of contact, 233, 258, 267
Angular momentum, 10
— — conservation of, 11
— — definition, 10
Anticlastic surfaces, 236
Antonow's rule, 259
Applications of gravity surveys, 65
Arrhenius, 5
Associated liquids, 265
Association in solution, 216
Astronomical evidence against any effect of temperature on gravitation, 46
— — for Newton's laws, 8
Atomic theory of matter, 1
Attracting rigid spheres model of liquid, 262
Axial vector, 8, 35

BASHFORTH'S and Adams' tables, 243
Beam, bending of, 87
Bending method for determining Young's modulus, 96
— moment, 88
— of beam, 87
Bennett, 266
Bernoulli, 3
Bernoulli's equation, 145
Berthelot, 185
Bifilar suspension, 41
Biharmonic functions, 85
Black soap films, 248
Boiling point, relation with osmotic pressure, 5, 211
Boltzmann's expression, 118
Borda's pendulum, 54
Boundaries, free and rigid. Effect on hydrodynamic problems, 159
Boundary layer, 4, 153, 159
Bowden, 273
Boys' experiments with soap films, 240
— measurement of G, 53
Breaking stress, 77
Bridgman's measurements at high pressures, 78, 199
Briggs (L. J.), 185
Brown and Harkins' relation, 255
Brownian movement, 120
Bubble, oscillating, 160
— soap, pressure inside, 236
Buckling load, 78, 103
Bulk modulus, 81, 100, 198
— viscosity, 139, 206
Butterworth, 160

CALLENDAR, 233
Cantilever, 93
Capillary rise method of measuring surface tension, 243, 251
Castor oil as lubricant, 274
Catalysts, 267
— poisoning of, 268
Catenoid soap films, 237
Cavendish's measurement of G, 52
Cavitation, 186
Cell model of liquid, 196
Central forces, 8, 12, 13
Chapman, 1
Circulation, 158
Clamped beams, 87
Clausius, 1
Cockcroft, 272
Coincidence method of timing, 55
Coker, 3
Colloids, 216
Combination tones, 291
Comparison of indices in dimensional analysis, 279
Compass, gyro-, 40
Compressional waves, 107
Compressive stress, 74
Conductivity, thermal, of gas, 121, 125
Conservation of energy, 10, 23
— — momentum, 9
Conservative forces, 12, 13
Contact, angle of, 233, 258, 267
Continuity equation of hydrodynamics, 141

Continuum theory of matter, 1
Cornu's method for elastic constants, 97
Cosmology, 70
Coulomb's theory of friction, 272
Critical adsorption, 272
— damping, 288
— region of liquid, 191
Crooke's radiometer, 135
Cylindrical soap films, stability of, 239

DALTON, 1
Damped oscillator, 285
Debye, 5, 197, 208, 223
Debye-Hückel theory of strong electrolytes, 223
Decrement, logarithmic, 287
Deep-water waves, 171
Deffrages, 61
Deflection of light-ray by heavy body, 69
Deformation by torsion, 86
— of beam due to bending moment and shear, 90
— — cantilever, 93
— — flat spiral spring, 94
— — solids, 72
Diameter, molecular, 123, 315
Diffusion, differential equation for, 228, 298
— in gases, 121, 126, 128
— — liquids, 227
— pumps, 129
— thermal, 127
Dilatation, 73
Dimensional analysis, 276
— constant, 283
Dimensionless quantity, 279
Dirac, 6
Dislocation, 102
Distribution function of liquid, 182
Double suspension mirror, 41, 52
Drift velocity in gases, 122, 140
Drop, dimensions of, 253
— oscillations of, 255
— pressure inside, 236
— stability of, 244
— weight of, 254
Dushman molecular gauge, 133
Dynamic similarity, 280, 282

EARTHQUAKES, 107, 111
Eckart, 207
Eddies, 4, 152
Einstein, 68, 69
Elastic constants, definitions of, 81
— — relations between, 82
— limit, 76
— stability, 103
Elastic waves, 107
Elasticity theorems, 78
Electrolytes, 220
— Debye-Hückel theory of, 223
— strong and weak, 223
Elevated masses, effect on gravity, 66
Ellipsoid, strain, 74
— stress, 75
Energy, conservation of, 10, 23
Eötvös torsion balance, 62
Equation of state of liquid, 186, 189
Equipartition of energy, 116
Euler, 3, 103, 142
Eulerian hydrodynamics, 142
Euler's criterion for buckling, 103
Ewing's extensometer, 95
Excluded volume of liquid, 190
Expansion viscosity, 139
Eyring's theory of holes, 197

FACTOR of safety, 78
Failure of Newton's laws, 68, 202
Faraday's theory of electrolysis, 220
Fatty acids, 248
— oils, 274
Ficks' law, 128, 229, 298
Films, adsorbed, solid, liquid, and gaseous, 257
— soap, 247
— unimolecular, 257
Finite arc of swing, effect on pendulum, 28, 56
Flat spiral spring, deformation of, 94
Flexural waves, 109
Floating bodies, forces between, 245
Force, central, 8, 12
— conservative, 12
— definition of, 7
— internal and external, 8
— on sphere in liquid, 278
— parallelogram law, 8
Forced oscillations, 289
Forces between floating bodies, 245
Fountain effect in liquid helium, 203
Fourier series, 291
— coefficients, 293
Fowler (R. H.), 226, 264, 269, 271
Franck, 123
Free oscillations, 285, 289
— path, mean, in gases, 121
— surface in hydrodynamics, 159
— — of rotating liquid, 146
Freezing point of solution, relation with osmotic pressure, 5, 213
Fresnel, 3
Friction, theories of, 272
Froude's number, 282
Fürth, 197

GALILEO, 2
Gas-like models of a liquid, 187
Gases, kinetic theory of, Chapter V, 113
Gauss, 1, 114
Gravitation, 2, Chapter III, 45
— at sea, 65
— Einstein's theory of, 68
Gravitational constant, measurement of, 50
— field, 46
— mass, 7, 49
— potential, 46
Gravity, acceleration of, 50, 54
— surveys, 62
— — applications of, 65
Gruen, 244
Guggenheim, 226
Gyro-compass, 40
Gyroscopic motion, 35

HALF-RANGE Fourier series, 294
Hamilton (W. R.), 2
Hammick, 263
Hardy, 259, 273
Harkins, 255
Harmonic oscillator, 18, 26, 285
Hawksbee's law, 145
Heat conductivity of gas, 125
Helium, liquid, properties of, 202
Henry's law, 187
Hertz, 123
Heyl, 54
Hittorf, 222
Hole model of liquid, 197
Hooke, 3
Hooke's law, 75
Hückel, 5, 223
Hugoniot, 177
Hydraulic jump, 179
Hydrodynamic virtual mass, 60, 161, 164
Hydrodynamics, 3, Chapter III, 138
Hydrostatics, 144

INCOMPRESSIBLE approximation in hydrodynamics, 143, 157
Increase of mass with velocity, 69
Inertial mass, 7, 49
Instantaneous axis of rotation, 33, 57
Interfaces, liquid, 257
Internal forces, 10
Inverse-square law, 3, 45
Ionic theory, 5, 220
Irrotational flow, 154, 158
Isotopes, separation of, by thermal diffusion, 127

JAEGER'S method of measuring surface tension, 251
Jeans, 124, 158
Jets, stability of, 155
Joly, 67
Joule-Thomson experiment, 178

KATER'S pendulum, 55
Kelvin, 3, 41, 52, 152, 229, 244
— double suspension mirror, 41, 52
Kinetic theory of gases, 1, Chapter V, 113
— — — — integrals occurring in, 300
Kittel, 208
Knudsen, 130, 135, 150
Knudsen's absolute manometer, 135

LAGRANGE, 2, 3, 142
Lagrangian hydrodynamics, 142
La Mer, 244
Lamb, 158, 160
Lamé, 3
Landenburg, 201
Langmuir, 268
Laplace, 2, 5, 231, 260
— theory of surface tension, 260
Latitude effect on acceleration of gravity, 66
Lennard-Jones, 1, 195, 197
Letter balance, 16
Leucohippus, 1
Liebermann, 207
Liquid films, 257
— helium, properties of, 202
Liquids, 1, Chapter VII, 182. See also Chapter VI, 138, Chapter VIII, 210, and Chapter IX, 231
— models of, 187, 196, 197
Logarithmic decrement, 287
Love, 3, 84
Lowering of vapour pressure, 212
Low-pressure phenomena in gases, 130

MACLEOD, 264
Manometer, Dushman, 133
— Knudsen's absolute; 135
— quartz fibre, 133
Mass, gravitational, 49
— inertial, 7, 49
Matthews, 265
Maxwell, 1, 114, 130, 200, 207
Maxwell's distribution law, 114
Mayer (J. E.), 233
Mean free path in gas, 121
Mechanics, 2, Chapter II, 7
Meinesz, 59
Millikan, 132

Mitchell, 266
Model experiments, 280
Moduli of elasticity, relations between, 82
Modulus of elasticity, 81
— — — Young's, 81
— — — rigidity or torsion, 81
— — — bulk, 81
Molecular diameters, 123, 315
Momentum, angular, conservation of, 10
— linear, conservation of, 9
Motion, Newton's laws of, 7
Mountains, effect of, on gravity, 50, 67

NEUMANN'S triangle, 258
Newton, 2, 7, 68, 154
Newton's laws, Chapter II, 7
Newtonian fluids, 140, 174
Non-linear hydrodynamics, 174
— — oscillator, 290
Non-Newtonian fluids, 140

ONE-DIMENSIONAL flow of fluid, 144 175
Organ pipe, 296
Orthogonal sets of functions, 292
Oscillations of drop, 255
Oscillator, overdamped and underdamped, 286
— simple harmonic, 18, 26, 285
Osmotic pressure, 5, 210
— — relations with vapour pressure, freezing and boiling points, 211
Ostwald's dilution law, 221
Overdamped oscillator, 286

PARACHOR, 4, 264
Parallelogram law of addition, 8
Pendulum, Borda's, 54, 277
— compound, 29
— Kater's, 55
— simple, 27, 277
Perrin, 120
Pfeffer pot, 211
Photo-elasticity, 3
Plasticity, 3, 77, 101
Plateau's spherule, 241
Pockels, 256
Poiseuille, 147, 150
Poiseuille's law, 147
Poisoning of catalysts, 268
Poisson, 3, 81, 98
Poisson's ratio, 81, 98
Polar vectors, 8
Potential, gravitational, 46
Potential, velocity, 167
Poynting's experiment, 51
Prandtl, 4
Probability theory, 114

QUARTZ fibre manometer, 133
Quincke, 4, 253, 254

RADIOMETER, 134
Rankine-Hugoniot equations, 177
Raoult's law, 217
Rational systems of units, 284
Rayleigh, 4, 165, 243, 254, 256
— disc, 165
Regnault, 198
Relativity theory, 2, 68
Repsold's pendulum, 60
Resistance of ships, 281
Resonance, 290
Reynolds, 4, 127, 153, 273, 282
Richards, 265
Riemann, 175
Riemann's function, 175
Rigidity modulus, 81, 98, 99, 100
Ripples, 170, 256
Roberts (J. K.), 267, 270
Rolling body, 34, 50
— disc, 38
Rotating liquid, 146
Routh's rule, 31, 95

SEARLE (G. F. C.), 95, 100
Second sound in liquid helium, 204
Sedimentation, 118
Seizing of metals, 267
Separation of isotopes, 128
Shafts, whirling of, 106
Shear, 73
— waves, 108
Shearing force, 88
Ships, resistance of, 281
Shock-front, 176
Similarity, dynamic, 280, 282
Simple harmonic oscillator, 18, 26, 285
— pendulum, 27, 277
Size of atoms, 120, 315
Solid adsorbed films, 257
—sphere, potential of, 49
Solids, adsorption on, 267
— deformation of, 3, Chapter IV, 72
— surface tension of, 266
Solubilities, 4, 186, 219
Solutions, 4, 210, 216
Southwell, 84, 106
Sphere in liquid, 161, 278
Spherical shell, potential of, 48

Spiral spring, deformation of, 99
Spreading of liquids, 258
Stability of cylindrical films, 239
— — struts, 103
Stokes, 3, 200, 206
Strain, 72
— ellipsoid, 74
Stream function, 167
— velocity, 140
Stress, 74
— calculations, 85
— ellipsoid, 75
Struts, stability and buckling load of, 103
St Venant, 3, 87
Sugden, 4, 264
Superfluidity, 202
Superheating of liquids, 245
Supersaturation, 193
Supported beams, 87
Surface, curved, effect on vapour pressure, 243
— energy, 232, 266
— tension, 4, 183, Chapter IX, 231
Sutherland, 124
Synclastic surfaces, 236

TAIT, 122, 199
Taylor (G. I.), 4, 152
Temperature, kinetic theory definition of, 115
Tensile strength of liquids, 185, 194, 245
— — — solids, 77
Tension, 74
Thermal conduction of gases, 121
— diffusion of gases, 121, 127
— transpiration of gases, 121, 127
Thermo-mechanical effects, 203
Thin soap films, 249
Thixotropy, 140
Tidal waves, 171, 173
Tides, 172
Topographical correction to gravity data, 66
Torpedo, steering of, 40
Torsion, 86
— modulus (*see* Rigidity modulus)
Tracer techniques, 128
Trouton's rule, 265
Turbulence, 4, 151
Two-fluid picture of liquid helium II, 203

ULTRASONIC absorption in liquids, 206
Underdamped oscillator, 286
Uniform strain, 72
Ursell, 195

VAN DER WAALS, 4, 189, 191, 260
Van't Hoff, 5, 214
Vapour pressure, over curved surface, 243
— — relation with osmotic pressure, 210
Velocity potential, 167
Virtual mass, hydrodynamic, 161, 164
— work, 15
Viscosity, 121, 124, 138, 147, 183, 199, 207
Von Jolly, 51
Von Karman, 152
Vortex-street, 152

WAVE equation, 107, 109, 156, 296
Waves on water, 166
West, 127
Wettability, 234
Whiddington, 97
Whirling of shafts, 106
Wilhelmy's method of measuring surface tension, 252
Wilson Chamber, 193
Work-hardening, 77, 101

XYLOPHONE, 110

YIELD of support, effect on pendulum, 58
Young (Thomas), 231, 260, 263
Young's modulus, 81, 90, 96, 97, 99

ZAHRADNICEK, 53

ALLAMA IQBAL LIBRARY
12031

JAMMU & KASHMIR UNIVERSITY
LIBRARY
No. 12031
Date. 19-5-55
SRINAGAR

PRINTED IN GREAT BRITAIN BY UNIVERSITY TUTORIAL PRESS LTD. FOXTON
NEAR CAMBRIDGE